ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Schmitz, Luca

Working Paper

Behind the curve: Reassessing inflation risks with a High-dimensional Distributional VAR (HiDVAR)

IMFS Working Paper Series, No. 245

Provided in Cooperation with:
Institute for Monetary and Financial Stability (IMFS), Goethe University Frankfurt am Main

Suggested Citation: Schmitz, Luca (2026) : Behind the curve: Reassessing inflation risks with a High-dimensional Distributional VAR (HiDVAR), IMFS Working Paper Series, No. 245, Goethe University Frankfurt, Institute for Monetary and Financial Stability (IMFS), Frankfurt a. M.

This Version is available at:
https://hdl.handle.net/10419/344237

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

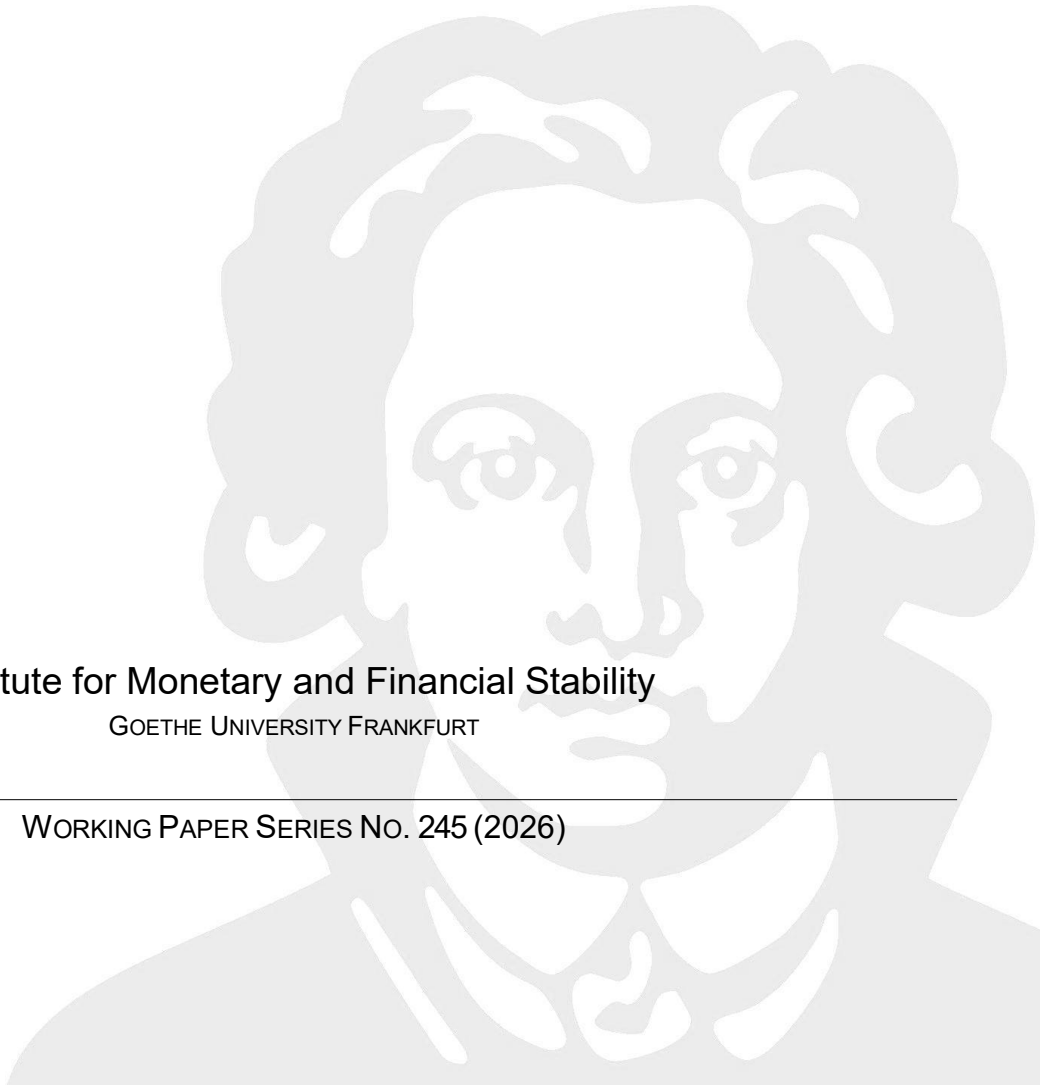
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Luca Schmitz

Behind the Curve: Reassessing Inflation Risks with a High-Dimensional Distributional VAR (HiDVAR)



Institute for Monetary and Financial Stability
GOETHE UNIVERSITY FRANKFURT

WORKING PAPER SERIES NO. 245 (2026)

This Working Paper is issued under the auspices of the Institute for Monetary and Financial Stability (IMFS). Any opinions expressed here are those of the author(s) and not those of the IMFS. Research disseminated by the IMFS may include views on policy, but the IMFS itself takes no institutional policy positions.

The IMFS aims at raising public awareness of the importance of monetary and financial stability. Its main objective is the implementation of the “Project Monetary and Financial Stability” that is supported by the Foundation of Monetary and Financial Stability. The foundation was established on January 1, 2002 by federal law. Its endowment funds come from the sale of 1 DM gold coins in 2001 issued at the occasion of the euro cash introduction in memory of the D-Mark.

The IMFS Working Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Institute for Monetary and Financial Stability
Goethe University Frankfurt
House of Finance
Theodor-W.-Adorno-Platz 3
D-60629 Frankfurt am Main
www.imfs-frankfurt.de | info@imfs-frankfurt.de



Behind the Curve: Reassessing Inflation Risks with a High-Dimensional Distributional VAR (HiDVAR)

Luca Schmitz

Institute for Monetary and Financial Stability (IMFS)

Goethe University Frankfurt

schmitz@imfs-frankfurt.de

www.lucaschmitz.net

September 2026

Keywords: Inflation, Forecasting and Simulation: Models and Applications, Central Banks and their policies, Nonparametric methods, Time-Series Models, Forecasting, Large Data Sets

JEL Codes: E31, E37, E58, C14, C32, C53, C55

Abstract

This paper develops a High-Dimensional Distributional Vector Autoregression (HiDVAR), a real-time, data-rich framework for forecasting the joint predictive distribution of GDP growth, inflation, and unemployment without imposing a parametric density shape. The motivation is policy risk monitoring: standard quantile-to-density pipelines can compress one-sided inflation risks in turbulent periods, especially when they rely on restrictive distributional modeling or forecast anchors. HiDVAR combines distribution regression (many binary threshold models), local projections, elastic-net shrinkage, and factor-based information extraction from an augmented real-time FRED-QD panel, and then constructs coherent multivariate, path-consistent scenarios using a triangular factorization and copula-based simulation.

In real-time out-of-sample evaluation (February 2013–June 2025), HiDVAR improves probabilistic forecast accuracy relative to the FRBNY Outlook-at-Risk framework and standard benchmarks, with the largest gains concentrated in inflation tail-risk assessment. Relative to the FRBNY model, HiDVAR reduces CPI inflation CRPS by 28% and the multivariate Energy Score by 22%, with a weakly significant Energy Score improvement in Diebold–Mariano tests. During the post-pandemic inflation surge, HiDVAR assigns substantially more probability mass to persistent high-inflation outcomes and produces earlier, better-calibrated upper-tail risk signals. A policy application maps the predictive distributions into risk-scenario probabilities and rule-implied policy-rate distributions, illustrating the value of flexible density forecasting for macro risk monitoring.

Contents

1 Introduction **1**

2 Literature **5**

3 Data **9**

4 Model **11**

5 Results **15**
5.1 Benchmarks 15
5.2 Calibration 15
5.3 Distributional Output 18
5.4 Out-of-sample Evaluation 19
5.5 Ablation of Model Components 25
5.6 Bandwidth Choice and Sensitivity 26

6 Risk Monitoring and Policy Translation **30**
6.1 Macro Risk Dashboard 30
6.2 Policy-Stance Implications 33

7 Conclusion **35**

A Distributional Scoring Rules and Statistical Tests **43**

B Parameters and restrictions used in the estimation **50**

C Copula Specification **51**

D Benchmarks **53**

E Stationarity of Variables **60**

F Robustness check with a different triangular ordering **61**

G Data **62**

1 Introduction

During the aftermath of the COVID-19 pandemic, Europe and the United States saw the largest surge in inflation since the early 1980s. Yet, most economists on both sides of the ocean did not see the surge coming or severely underestimated it (Koch & Noureldin, 2024). One possible explanation is that prevailing risk assessments did not adequately capture distributional uncertainty. Probabilistic forecasts are particularly valuable during periods of economic stress for analyzing the risks of, for example, high inflation or recession probabilities. However, many standard forecasting models rely on symmetric, often normal, distributions for their projections, including most Bayesian Vector Autoregressions (BVARs), Dynamic Stochastic General Equilibrium (DSGE) models, or simple local projections according to Jordà (2005). This can severely underestimate one-sided or asymmetric risks. An example of this can be seen in the Outlook-at-Risk (OaR) approach by the Federal Reserve Bank of New York (FRBNY) staff[1]. OaR provides the distributional 1-year-ahead forecasts for output, inflation, and unemployment in the US. They construct quantile regressions based on forecasting errors of the Blue-Chip Consensus Forecast by Wolters Kluwer, which they then interpolate to densities (Federal Reserve Bank of New York, 2023). The model is largely based on the Growth-at-Risk model by Adams, Adrian, Boyarchenko, and Giannone (2021). This method, while not imposing symmetry by using skew-t distributions, nevertheless can result in symmetric, unimodal, generic distributions. Furthermore, such a model is heavily dependent on the validity of professional forecasts, which have been notably poor during the inflation surge (Gourio & Sarma, 2025). To show this shortcoming, I group 1-year-ahead forecasts by various institutions and sources for the growth of Personal Consumption Expenditure (PCE) prices. These include the Federal Open Market Committee (FOMC), the Congressional Budget Office (CBO), the Survey of Professional Forecasters (SPF), the Business Survey of the Federal Reserve Bank of Atlanta, the Federal National Mortgage Association (Fannie Mae), and Household inflation expectations by the FRBNY[2].

The results are underwhelming. Despite the peak of YoY PCE at around 7% in 2022Q2, no institutional forecast ever reaches 4%. Only when inflation normalized in 2023 and 2024 did the forecasts improve. It seems like most institutions were too focused on mean reversion to the 2% target in the medium run. The SPF and CBO forecasts especially undershot inflation: at no point did either project 1-year-ahead PCE inflation exceeding 2.7%. Only the households did somewhat anticipate prolonged inflation but also tended to predict highly persistent inflation even during the disinflation periods, making the forecasts similarly uninformative. For CPI, where not as many forecasts are available, the story is a similar one, with a 1-year-ahead

[1] "OaR is a staff research product published monthly on the New York Fed's website; per the Bank's own disclaimer, it 'is not an official forecast of the Federal Reserve Bank of New York, its President, the Federal Reserve System, or the Federal Open Market Committee' Federal Reserve Bank of New York (2023). For brevity, we follow common usage and refer to it as 'FRBNY' or 'the FRBNY model' throughout, denoting Outlook-at-Risk specifically.

[2] As both the CBO and the FOMC meet in irregular intervals, there are some missing values. Since the FOMC doesn't provide explicit YoY forecasts but rather annual inflation projections, I construct a weighted average based on the timing. For example, if the meeting is held in September 2021, I would get the YoY forecast by taking $\frac{1}{4}$ of the 2021 forecast and $\frac{3}{4}$ of the 2022 forecast since September is the ninth month.

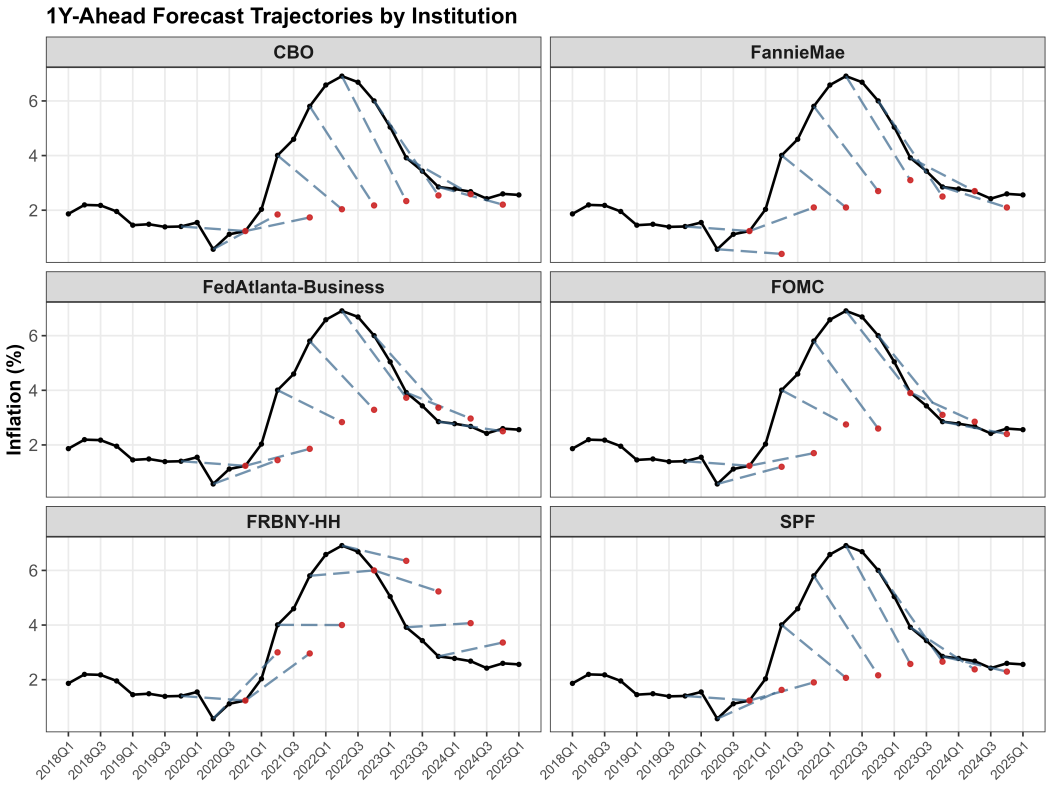


Figure 1: 1-year ahead Inflation forecasts against actual values

forecast maximum of the SPF mean at 3.3%, with CPI rising up to 9% in reality. Generally speaking, the implied pull of mean reversion and the central bank target of 2% are quite surprising.

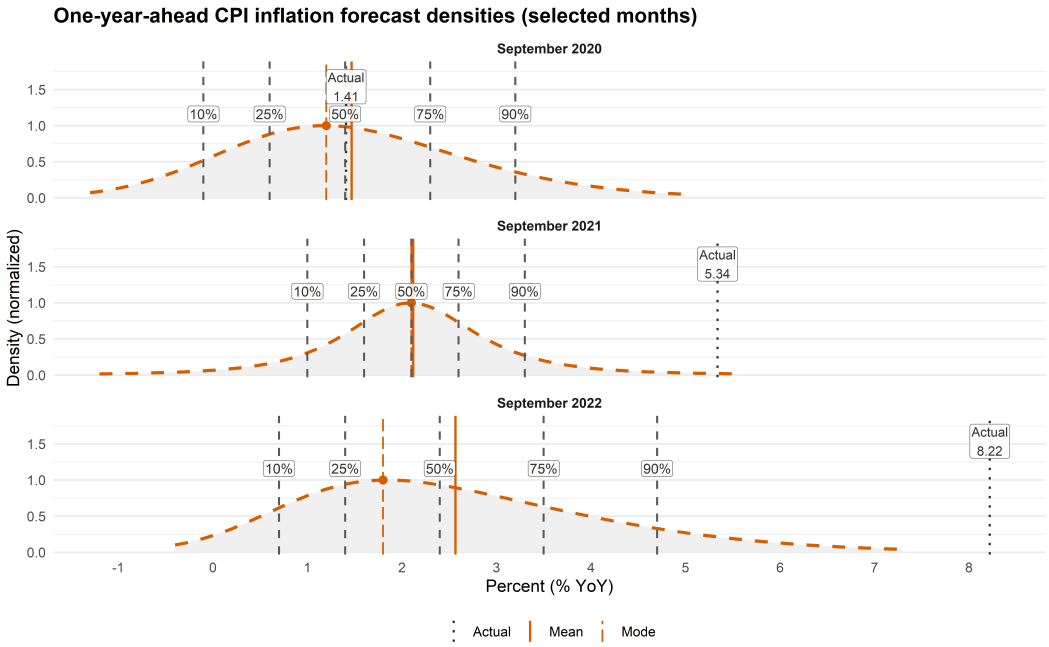
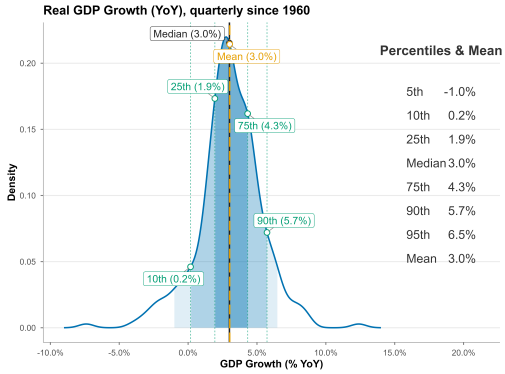


Figure 2: FRBNY OaR 1-year-ahead CPI Forecasts; Origins: September 2020, 2021, and 2022; actual marks the inflation value at the forecast origin

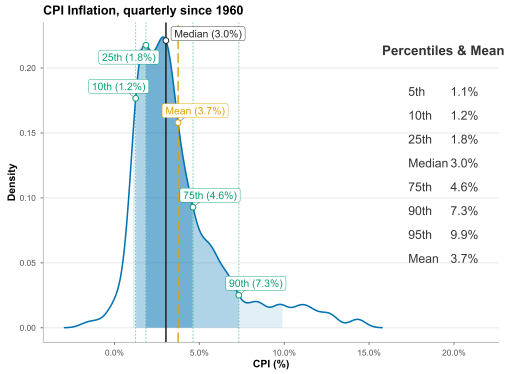
Such an analysis makes it difficult to justify relying on professional forecasts as the main data

source. As an illustration of the weaknesses, Figure 2 shows the September 2021 densities produced by the OaR model by the FRBNY, targeted at September 2022. Especially the CPI inflation forecast is concerning. In September 2021, where YoY inflation already reached 5.34%, predicting an almost symmetric and unimodal distribution centered around 2.1% appears like a debatable risk forecast. Also, putting roughly the same risk on deflation and inflation rising above 4% is debatable, given the prevalent inflation.

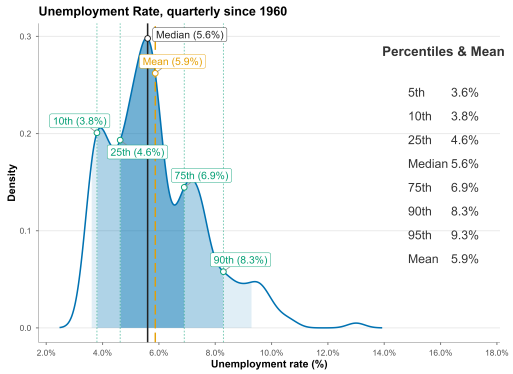
To highlight the distance to inflation realizations of such forecasts, I take a look at the empirical distributions of the variables. For that, I analyze empirical distributions of their targets in the US since 1960. While GDP growth indeed has a fairly standard distribution, and the unemployment rate has somewhat symmetric quantiles, the story is different for inflation. For CPI, there are barely any episodes of deflation but significant examples of high inflation. Hence, a symmetric prediction for inflation in recent times seems not well-suited given historic trends. Furthermore, the common assumption of unimodality in risk forecasts is debatable from this as well.



(a) GDP growth



(b) CPI inflation



(c) Unemployment rate

Figure 3: Empirical Year-on-Year Quarterly distributions from 1960–2025 with selected quantiles

To address these limitations, I built on the distributional VAR framework of Wang, Oka, and Zhu (2026). The great advantage of this approach is that it does not rely on any direct distributional assumptions like picking a fixed distribution family but rather lets the data decide on the full shape of the CDF. I extend their model by including machine learning techniques in a large real-time dataset and by allowing for path-consistent draws, using copulas. Data from an augmented version of the FRED-QD (McCracken & Ng, 2021) is used by extracting factors and utilizing optimal elastic net penalization. The out-of-sample period stretches from February 2013 to June 2025 with monthly 4Q-ahead projections. Results include a

notable shift in inflation distributions during the inflation surge, with much more prevalent tail risks. The model also delivers substantial improvements in probabilistic forecast performance, particularly for inflation, as well as the possibility of creating a macro risk dashboard with predictive scenario probabilities and distributional prescriptions for the policy rate.

Going forward, section 2 summarizes relevant literature, while section 3 dives into the augmented FRED-QD dataset used. Section 4 describes the model setup and estimation, sections 5 and 6 present the results and policy applications, respectively, and section 7 concludes.

2 Literature

Central banks and policy institutions increasingly communicate not only point forecasts, but also the uncertainty and balance of risks surrounding their projections. An early and influential example is the Bank of England's inflation fan chart, which represents the forecast distribution using a parametric two-piece distribution (Britton, Fisher, & Whitley, 1998), often called the 'River of Blood'. Subsequent work examines both the statistical interpretation and the calibration of such fan charts (Clements, 2004; Wallis, 1999). This literature established two principles that motivate this paper. First, the conditional mean or modal forecast is generally insufficient when the predictive distribution is asymmetric, subject to substantial tail risk or exhibits pronounced uncertainty. Second, the usefulness of a density forecast depends not only on how it is communicated, but also on whether its probability assignments are calibrated.

The policy motivation is closely related to the risk-management view of monetary policy. When policymakers face asymmetric losses, the probability of particularly adverse outcomes can matter independently of the central projection (Kilian & Manganelli, 2008). This distinction became especially salient during the post-pandemic inflation surge. Once inflation had moved materially above target, the relevant policy question was no longer simply whether the conditional mean would decline. It was also how much probability should be assigned to inflation remaining persistently high, and whether this risk occurred jointly with weakening output or labor-market conditions. Reliable policy risk monitoring therefore requires predictive distributions whose location, dispersion, skewness, tail thickness, and potentially modality can change with economic conditions.

A major modern strand of the literature estimates such risks using conditional quantiles. The Growth-at-Risk framework maps financial conditions into the lower tail of future GDP growth using quantile regressions (Adrian, Boyarchenko, & Giannone, 2019). Adams et al. (2021) extend this architecture into a broader macroeconomic forecasting system by estimating conditional quantiles of forecast errors around professional forecast anchors. The operational Outlook-at-Risk (OaR) system, which serves as the main policy benchmark for this paper, applies this approach to GDP growth, inflation, and unemployment: conditional quantile regressions, based on the Composite Indicator of Systemic Stress (CISS), produced by the ECB and Consensus Forecast errors, are estimated on a finite probability grid, after which a parametric distribution is fitted to obtain a complete predictive density (Federal Reserve Bank of New York, 2023). Related work shows that macroeconomic predictive distributions can display substantial state dependence and, in particular, that multimodality can be economically important (Adrian, Boyarchenko, & Giannone, 2021).

A more specialized inflation-forecasting literature studies the conditional distribution of inflation directly. Manzan and Zerom (2013) show that macroeconomic predictors contain information about different parts of the US inflation distribution. Manzan (2015) extends distributional forecasting to a data-rich environment, while Korobilis (2017) combines inflation quantile regressions with Bayesian model averaging using real-time US data. More recently, López-Salido and Loria (2024) formalize the concept of Inflation-at-Risk and document that substantial variation in inflation tails can remain hidden when attention is restricted

to the conditional mean. For the euro area, Lenza, Moutachaker, and Paredes (2025) use quantile regression forests and a broad set of inflation determinants to construct nonlinear density forecasts. Grishchenko and Wilcox (2024) instead combine a state-space framework with survey information to study inflation tails. Finally, Schröder (2026) develops a flexible density-regression framework that decomposes movements in the full inflation distribution into contributions from the underlying predictors. Together, these contributions demonstrate that inflation risk is time varying, can be driven by different variables at different parts of the distribution, and need not be summarized adequately by a symmetric forecast interval around a point prediction.

The quantile-to-density step nevertheless introduces an additional modeling choice. A finite collection of estimated quantiles does not by itself define a smooth, complete predictive density. The missing parts of the distribution must be obtained through interpolation, smoothing, or the fitting of a parametric family. A skew-t distribution, for example, can accommodate asymmetry and heavy tails, but remains restricted to a particular, typically unimodal, functional form. Consequently, irregularities in the estimated quantiles, including shoulders, flat regions, or multiple modes, may be attenuated during density construction. Mitchell, Poon, and Zhu (2024) study this problem directly and show that nonparametric quantile-to-density construction can reveal multimodality that is obscured by a fitted parametric density. Parametric smoothing is therefore not inherently inappropriate, it can stabilize noisy quantile estimates and facilitate communication, but its empirical adequacy must be evaluated rather than assumed.

A related issue concerns the use of an external point-forecast anchor. Professional forecasts may incorporate judgment and information not captured by an econometric model, and anchoring can stabilize the location of the resulting predictive distribution. At the same time, persistent errors in the anchor are transmitted to the density unless the estimated conditional forecast-error distribution adjusts sufficiently to offset them. These considerations motivate a complementary approach that estimates the conditional distribution directly from the available macroeconomic information, without centering it mechanically on an external consensus forecast and without pre-specifying a parametric family for the marginal predictive density.

The statistical foundations for such an approach are provided by quantile and distribution regression. Classical quantile regression estimates selected conditional quantiles without specifying the complete conditional distribution (Koenker & Bassett, 1978), while methods such as quantile regression forests allow the quantiles to depend nonlinearly on a large predictor set (Meinshausen, 2006). Distribution regression instead estimates the conditional cumulative distribution function directly by fitting binary regressions for indicators of the form $\mathbb{1}\{Y \leq c\}$ over a grid of thresholds c (Foresi & Peracchi, 1995). Because each threshold or quantile can be estimated separately, finite-sample estimates need not be monotone. Monotone rearrangement restores the ordering required for a valid CDF or a non-crossing quantile function and weakly improves the approximation under common loss metrics (Chernozhukov, Fernández-Val, & Galichon, 2010). Distribution regression is particularly attractive here because it targets the CDF itself and therefore avoids an additional parametric quantile-to-density fitting step.

A growing literature embeds conditional quantiles or distribution regressions in multivariate time-series systems. Quantile VARs allow the propagation of shocks and predictive relationships to differ across parts of the conditional distribution. Surprenant (2025) provides a large forecasting evaluation of QVARs for US macroeconomic variables and shows that their performance can differ materially from Gaussian, GARCH, and stochastic-volatility VAR benchmarks. Korobilis and Schröder (2025) develop a quantile factor-augmented VAR for

monitoring macroeconomic risks across euro-area countries, thereby combining multivariate quantile dynamics with high-dimensional factor information. These models are closely related to this paper, but primarily organize the system around a selected collection of conditional quantiles.

The closest methodological antecedent is the distributional VAR of Wang et al. (2026). The DVAR estimates a sequence of conditional distribution regressions and uses a recursive triangular factorization to assemble them into a coherent multivariate conditional distribution. Its principal advantage is that the marginal conditional distributions are not confined to a Gaussian, skew-t, or other fixed density family. The approach can therefore accommodate state-dependent asymmetry, changing tail thickness, and multimodality, subject to the maintained regression specification, the selected threshold grid, the recursive ordering, and monotonicity. Recent related work uses joint distributional forecasts to study growth–inflation tail configurations and stagflation risks (Corradi & Llorens-Terrazas, 2026). This literature moves beyond separate univariate at-risk measures by treating macroeconomic risk as an inherently joint object.

A further strand concerns the information set used to forecast these distributions. Dynamic factor models summarize large panels of macroeconomic and financial variables using a small number of common factors (Bai & Ng, 2002; Stock & Watson, 2002). They provide the foundation for data-rich forecasting and real-time nowcasting systems that process staggered data releases and ragged-edge information (Doz, Giannone, & Reichlin, 2011; Giannone, Reichlin, & Small, 2008). Machine-learning evidence similarly indicates that nonlinearities and regularization can be especially useful in unstable macroeconomic environments (Goulet Coulombe, Leroux, Stevanovic, & Surprenant, 2022). In large multivariate systems, Bayesian shrinkage (Bańbura, Giannone, & Reichlin, 2010; Giannone, Lenza, & Primiceri, 2015) and penalized regression offer complementary ways to control estimation variance as the number of predictors increases.

The real-time dimension is important in its own right. Forecasts constructed from revised historical data can use information that was unavailable at the forecast origin and can consequently provide an overly favorable assessment of predictive performance. The real-time forecasting literature therefore emphasizes vintage-consistent estimation and evaluation (Croushore & Stark, 2001; Koenig, Dolmas, & Piger, 2003). This consideration is particularly important during turning points and periods of rapid inflation change, when publication lags, data revisions, and the evolving information set can materially alter the apparent forecast problem. Large vintage databases such as FRED-QD provide a natural starting point for such exercises (McCracken & Ng, 2021), but their quarterly frequency also motivates an explicit nowcasting layer that incorporates more timely monthly information.

Policy applications additionally require forecasts that are coherent across variables and over time. A recession, stagflation episode, or policy-rate path cannot generally be evaluated from independently drawn one-period marginal forecasts. Copulas provide a way to separate marginal distribution estimation from dependence modeling (Sklar, 1959). In time-series applications, structured Gaussian or related copulas offer a tractable means of mapping flexible marginal distributions into joint simulations (Patton, 2006). For the present application, the recursive distributional system provides conditional dependence among variables within each forecast horizon, while a structured copula layer organizes the probability transforms across variables and horizons. The resulting draws can be aggregated into year-on-year outcomes and classified into internally consistent macroeconomic scenarios.

Finally, density forecasts must be evaluated using criteria that reward both calibration and sharpness. Proper scoring rules provide the formal foundation for such comparisons (Gneiting & Raftery, 2007). The continuous ranked probability score evaluates the full univariate CDF and is less locally sensitive than the log score to very small predictive densities at the realization. Its multivariate analogue, the energy score, can be applied to joint forecast draws. Because the main policy question concerns upper-tail inflation risk, threshold- or tail-weighted proper scores provide an important complement to average full-distribution performance (Gneiting & Ranjan, 2011). PIT-based diagnostics assess marginal calibration (Diebold, Gunther, & Tay, 1998), while score-based tests can evaluate calibration of multivariate forecast distributions (Knüppel, Krüger, & Pohle, 2022). Decompositions of the mean CRPS further distinguish gains arising from lower miscalibration from those reflecting stronger discrimination across economic states (Arnold, Walz, Ziegel, & Gneiting, 2024).

This paper contributes at the intersection of inflation-risk forecasting, data-rich distributional regression, and multivariate macroeconomic risk analysis. Methodologically, it builds most directly on the DVAR framework of Wang et al. (2026), estimating recursively ordered, horizon-specific distribution regressions and imposing monotonicity through rearrangement (Chernozhukov, Fernández-Val, & Galichon, 2010). HiDVAR extends this framework in three main dimensions. First, it introduces a real-time, high-dimensional information set summarized by macroeconomic factors and combines it with elastic-net regularization and externally validated tuning. Second, it estimates the distribution regressions as horizon-specific local projections and embeds the resulting forecasts in a structured copula simulation layer, producing multivariate paths that are coherent across the four-quarter forecast horizon. Third, it evaluates the resulting distributions in a real-time out-of-sample exercise against institutional OaR densities and econometric and machine-learning benchmarks.

Relative to OaR and other quantile-to-density systems (Adams et al., 2021; Federal Reserve Bank of New York, 2023; Mitchell et al., 2024), HiDVAR does not center the predictive distribution on a professional consensus forecast and does not fit a parametric family to a finite set of predicted quantiles. Relative to the inflation-at-risk and quantile-regression-forest literatures (Lenza et al., 2025; López-Salido & Loria, 2024; Schröder, 2026), it jointly models inflation, output, and unemployment and produces serially coherent economic paths rather than only marginal inflation distributions. Relative to QVAR and QFAVAR approaches (Korobilis & Schröder, 2025; Surprenant, 2025), it estimates the conditional CDF directly over a dense outcome grid rather than organizing the model around a limited collection of quantile levels. Finally, relative to the original DVAR, the principal contribution is not distribution regression itself, but its high-dimensional real-time implementation, prediction-oriented regularization, cross-horizon scenario generation, and application to inflation-risk monitoring. These extensions allow the joint predictive distribution to be translated into inflation-threshold probabilities, mutually exclusive macro-risk scenarios, and rule-implied policy-rate distributions, while being evaluated using marginal and multivariate scoring rules and calibration diagnostics.

3 Data

The main data source for this paper is the quarterly FRED-QD dataset (McCracken & Ng, 2021). It comprises around 248 variables[1] for the US dating back to 1959Q1. This dataset is commonly used in forecasting with factors due to its richness in both variables and time-series length. It includes variables from every important category like NIPA, Employment, Prices, Financial Markets, Housing, Interest Rates, and many more. This real-time dataset has been originally published monthly since May 2018.

However, as this would comprise a rather short, real-time, out-of-sample period, I built my own dataset mirroring FRED-QD, with a few additional variables, as well as a longer vintage history. Due to variable losses because of missing vintage series, as well as the availability of CISS predictors for the main benchmark by the FRBNY in real-time, February 2013 marks the start of the out-of-sample forecasting period[2]. To be more precise, in 2013, the Bureau of Economic Analysis (BEA) started including more granular GDP and price level data, explaining the large drop in available variables. I coin the new dataset FRED-QD+, in analogy to its heritage. I chose this dataset on purpose over the popular monthly counterpart FRED-MD (McCracken & Ng, 2016) due to the larger number of variables and the importance of GDP as an outcome variable here, which is only available long-term as a quarterly or yearly measure. However, due to the quarterly structure, FRED-QD(+) can be slow to react to, for example, changes in inflation, which is published monthly, and as the dataset does not include direct nowcasts, the May and June releases, for example, are still concerned with Q1 values as their latest observations.

To circumvent this, I build my own nowcasting layer based on current-quarter observations, as around two-thirds of variables are available at a monthly frequency. The nowcasting values stem from various sources, mainly directly from ALFRED and FRED-MD, where it is possible. For important variables like the two outcome variables GDP growth and CPI inflation, I directly use the real-time nowcasts provided by the Atlanta FED ('GDPNow') and the Cleveland FED (Inflation Nowcasting Project). Also, if available at the vintage date, nowcasts can be filled with the nowcasts of the Survey of Professional Forecasters (SPF). If a variable is available from more than one source, the institutional nowcasts are always preferred. Nowcasts are not created for the January, April, July, and October releases, as almost no representative current-quarter data would be available.

For the nowcasting layer, I split the dataset into monthly and quarterly variables to estimate a Mixed-Frequency Dynamic Factor Model, making use of current-quarter observations of the monthly variables. Essentially, a dynamic factor model (DFM) is a Kalman-filter-type model that uses PCA factors extracted from the dataset for the measurement equation. Now, a DFM is estimated on the monthly panel via Kalman filter smoothing, following Bańbura and Modugno (2014); Doz, Giannone, and Reichlin (2012). I first represent the data in a state-space form, where principal component–based factors provide initial estimates of the

[1] This number can change over time, as the database changes from time to time due to discontinued series.
[2] For more details on the exact data specifications, see G

latent state:
Measurement equation:

$$y_t = \Lambda f_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, R) \tag{1}$$

State-Space equation:

$$f_t = \sum_{p=1}^{P} \Phi_p f_{t-p} + u_t, \qquad u_t \sim \mathcal{N}(0, Q) \tag{2}$$

with y_t denoting the observed series (containing missing entries), f_t the latent factors, Λ the factor loading matrix, ε_t the idiosyncratic errors, u_t the factor innovations, and Φ_p denotes the factor transition matrices. Lags are chosen according to the Hannan-Quinn information criterion (Hannan & Quinn, 1979).

The Kalman filter uses the available observations to recursively update the factor estimates in real time, naturally handling periods where some series are missing by updating only on the available components. The Kalman smoother then refines these estimates over the entire sample, exploiting both past and future information. Finally, the smoothed factor estimates, together with the factor loadings, are used to reconstruct the missing observations as their model-implied conditional expectations given all available data.

Afterwards, the factors are aggregated to a quarterly base and regressed on the quarterly part of the dataset to bridge the DFM to a quarterly level. Using this, I can impute not only missing values within the dataset, but also at the ragged edges, including the nowcast, based on the already available monthly observations in said quarter[3]. Of course, this is only feasible when there are current-quarter observations, which is why this task is not executed for the first month of the quarter, for example, January or April. This method provides a balanced panel dataset I can use for this analysis. However, the dataset is too large to work with directly. Thus, the information is captured by factors. To extract factors, I rely on the methods of Bai and Ng (2002) to extract an endogenously determined number of factors. The outcome variables are, of course, excluded from the factor calculation.

To ensure a true real-time setup, each vintage is treated separately where all factor extractions, DFMs, and missing value imputations are executed only for the data available at that vintage to prevent foresight bias. Later data revisions or vintages do not interfere with the estimation of the vintage at use.

To capture past relationships, four lags of every variable are included. This leaves me with a dataset with dimensions TxN of at least 215x72[4]. To handle the still large number of regressors, elastic-net penalization with external-validation hyperparameter tuning inspired by Chi, Fan, Ghigliazza, Giannone, and Wang (2025) is applied. This two-step compression-penalization approach is similar to the work of Kim and Swanson (2014).

[3] For a more detailed derivation and explanation, consult Appendix G

[4] Of course, the time series dimension grows with new observations. In practice, there are often more than 100 explanatory variables, including lags, due to multiple factors per group.

4 Model

The model generalizes distributional regression to a multivariate setting with large data sets and serially dependent forecasts. It is largely based on the distributional VAR setup of Wang et al. (2026). I extend their framework by including penalized regressions with hyperparameter tuning with external validation and a copula construction to model the path-consistency given stationarity concerns, such that the full trajectory of the variables' forecasts can be evaluated. To honor the large dataset, the machine learning methods, and the works of Wang et al. (2026), the model is called a High-Dimensional Distributional VAR (HiDVAR). Of course, this name is debatable as it does not constitute a classical VAR. However, the SVAR-style ordering and the multivariate output give rise to a naming convention like this, as well as the DVAR naming in Wang et al. (2026).

At the heart of the forecasts are local projections, introduced by Jordà (2005). These estimate the impulse responses at each horizon h by running separate regressions of the form

$$y_{t+h} = \mu_h + \beta_h x_t + \gamma_h' Z_t + \varepsilon_{t+h}, \quad h = 0, 1, \ldots, H, \tag{3}$$

where

- y_{t+h} is the outcome h periods ahead,
- x_t is the shock or variable of interest,
- Z_t is a set of control variables (e.g. lags of y_t and x_t),
- β_h traces out the impulse response at horizon h.

This approach avoids iterative forecasting and is robust to model misspecification. When no shocks are applied, local projections reduce to horizon-h forecasts conditional on the exogenous regressors, which is the case considered in the main model.

To compare the model directly to the FRBNY approach, the model's outcome variables are GDP growth, CPI inflation, and the unemployment rate. For stationarity purposes, I use QoQ log changes for GDP and inflation, as well as first differences for the unemployment rate.

However, local projections in this form are univariate. To obtain full predictive distributions, the local projections are wrapped in distributional regressions, similar to Chernozhukov, Fernández-Val, and Galichon (2010). They show how binary-response regressions at many outcome thresholds can be stitched together into a full conditional distribution, without parametric distributional assumptions, allowing for multimodal and skewed shapes. Wang et al. (2026) extend this idea to a multivariate distributional regression.

Consider a univariate outcome Y and covariates Z. A convenient way to estimate the entire conditional distribution $F_{Y|Z}(y|z)$ is to fit a binary-response model at many outcome thresholds $y \in \mathcal{Y}$, which here are characterized by an extended grid of empirical quantiles of the target variable,

$$\widehat{F}_{Y|Z}(y \mid z) = \Lambda\big(\phi(z)^\top \widehat{\theta}(y)\big),$$

where Λ is a logit link and $\phi(z)$ collects basis terms in z. In a nutshell, the probability of $P(y_{t+h} < y_{cut}|z)$ is calculated for a large number of y_{cut}, such that they can be generalized into a Cumulative Distribution Function (CDF) without any pre-determined shapes. A detailed asymptotic analysis of the penalized DR estimator is beyond the scope of this paper.

Because each y is estimated separately, finite-sample noise can violate monotonicity in y. In this case, the grid points y are empirical quantiles of the outcome variables (see Figures 3a–3c) with additional padding of two standard deviations on both ends. Let $y_1 < \cdots < y_m$ be such a grid and form the vector $\big(\widehat{F}_{Y|Z}(y_j \mid z)\big)_{j=1}^{m}$. The increasing rearrangement sorts these values into nondecreasing order, producing $\widehat{F}^{\uparrow}_{Y|Z}(y_j \mid z)$ with $0 \le \widehat{F}^{\uparrow}_{Y|Z}(y_1 \mid z) \le \cdots \le \widehat{F}^{\uparrow}_{Y|Z}(y_m \mid z) \le 1$. This yields a proper, monotone conditional CDF, and (in L_p metrics) the rearranged curve is weakly closer to the target than the unrearranged fit. A conditional density can then be obtained by smoothing the derivative of $y \mapsto \widehat{F}^{\uparrow}_{Y|Z}(y \mid z)$, e.g. via finite differences or kernel differentiation, which preserves nonnegativity and unit integral up to numerical error.

Now let $Y = (Y_1, \ldots, Y_J)^\top$ and fix an ordering $1, \ldots, J$. Define the conditioning stacks

$$X_1 := (1, Z^\top)^\top, \qquad X_j := (1, Z^\top, Y_1, \ldots, Y_{j-1})^\top \quad (j = 2, \ldots, J),$$

and estimate each marginal conditional CDF $F_{Y_j|X_j}(y_j \mid x_j)$ by distribution regression across a grid in y_j, followed by increasing rearrangement in y_j to obtain $\widehat{F}^{\uparrow}_{Y_j|X_j}(y_j \mid x_j)$. The multivariate conditional distribution is then assembled by the triangular (Rosenblatt) factorization (Wang et al., 2026). For $J = 2$, the conditional joint CDF is

$$F_{(Y_1,Y_2)|Z}(y_1, y_2 \mid z) = \int_{-\infty}^{y_1} F_{Y_2|X_2}(y_2 \mid 1, z, s)\, dF_{Y_1|X_1}(s \mid 1, z),$$

and for a general J, one proceeds recursively by integrating $F_{Y_j|X_j}(\cdot \mid 1, z, y_1, \ldots, y_{j-1})$ with respect to the already constructed $F_{(Y_1,\ldots,Y_{j-1})|Z}(\cdot \mid z)$. Replacing each component by its rearranged estimator yields

$$\widehat{F}_{Y|Z}(y \mid z) := \rho\big(\widehat{F}^{\uparrow}_{Y_1|X_1}, \ldots, \widehat{F}^{\uparrow}_{Y_J|X_J}\big)(y \mid z),$$

which is coordinate wise nondecreasing in y and bounded in $[0, 1]$, hence a proper joint CDF. When densities exist, the corresponding joint conditional density inherits the triangular chain rule,

$$f_{Y|Z}(y \mid z) = f_{Y_1|X_1}(y_1 \mid 1, z) \prod_{j=2}^{J} f_{Y_j|X_j}\big(y_j \mid 1, z, y_1, \ldots, y_{j-1}\big),$$

where each factor can be obtained by differentiating the rearranged marginal CDF in its outcome coordinate. The ordering encodes a recursive identification (e.g., lower-triangular schemes familiar from VARs) and can be chosen to match structural assumptions; alternative orderings deliver alternative (still valid) factorizations of the same joint law. I employ a conditional factorization using a recursive (Cholesky) scheme with the ordering $(GDP \to UR \to CPI)$. The ordering reflects a within-quarter timing assumption commonly used in applied macro VARs: real activity can adjust contemporaneously to shocks through production, hours, and utilization, while labor-market quantities adjust with frictions and therefore respond with delay, and inflation is the most sluggish variable due to staggered price setting and contract structures. Accordingly, unemployment and inflation do not respond contemporaneously

to shocks ordered after them. I assess sensitivity by considering alternative orderings in Appendix F; the qualitative conclusions are unchanged. The fact that the ordering has little effect confirms that this copula-based dependence structure is successfully capturing the joint dynamics, making the results robust to this specific identification choice.

Extending beyond Chernozhukov, Fernández-Val, and Galichon (2010) and Wang et al. (2026), each binary-response local projection is estimated as a penalized logit model with an elastic-net penalty on the coefficients due to the large number of regressors. The elastic-net penalty takes the form $\lambda\{(1-\alpha)\|\theta\|_2^2/2 + \alpha\|\theta\|_1\}$, with $\alpha \in [0,1]$ and $\lambda > 0$. The regularization parameters (α, λ) are optimized in a data-driven way: for each variable and horizon they are tuned using a two-step, rolling external validation procedure in the spirit of Chi et al. (2025), but targeting the continuous ranked probability score (CRPS)[1] of the implied predictive distribution rather than pinball loss in the original paper. In practice, this translates to a 40-period (10 years) hyperparameter tuning on the training set before the forecasting exercise, here 2002Q4-2012Q4, in which the model forecasts with a grid of α values with tuned λ for all horizons. For robustness, this is augmented by an energy score [1] sanity check to confirm the selection and potentially decide between close contenders. The resulting hyperparameters can be found in Appendix B. The penalization introduces a standard shrinkage bias, but typically reduces variance and improves predictive performance, particularly when the number of lags and covariates is large. I therefore view the estimator as a prediction-oriented regularization of the distributional regression in Wang et al. (2026).

However, this still leaves a final issue. Since the YoY values of these variables are not stationary (see Appendix E for formal unit-root tests), I estimate distributional regressions for quarter-on-quarter (QoQ) changes instead. To be precise, QoQ growth rates for GDP and CPI, as well as first differences for the unemployment rate. This yields, for each horizon, a multivariate distribution for the QoQ outcomes, which are initially independent across horizons. To obtain path-dependent YoY forecasts, I need to add up the QoQ predictions while accounting for serial dependence. To this end, I use a Gaussian copula with a Kronecker structure to connect the horizon-specific distributions across time, such that all draws for the different horizons are taken jointly and are serially correlated[2]. This ensures path dependence: For example, a high inflation draw today makes a high draw tomorrow more probable. In the end, I obtain a Monte-Carlo-like sample of n different economic paths over the next four quarters, which I then aggregate into YoY forecasts and risk scenarios.

To summarize, instead of fitting an existing distribution to the data, the model uses probabilities from individual cutoffs to build a distribution from the ground up. A high-dimensional dataset is incorporated with a two-step tuned elastic net framework based on PCA factors. A triangular factorization, as well as a path-consistent copula setup, ensures multivariate transmission. The result comprises non-trivial, multivariate predictive distributions connected over time by the copula structure.

As this setup, especially with hyperparameter tuning, is computationally involved, it is estimated and simulated using the Linux-based FUCHS cluster by the Center for Scientific Computing (CSC) at Goethe University. FUCHS is a regional Tier 3 general-purpose CPU cluster. Due to heavy parallelization, running the full project including all periods without hyperparameter tuning takes only about four hours on the cluster. The hyperparameter tuning for the elastic net parameters α and λ takes around 36 hours, which, however, is only a

[1] A short explanation of both CRPS and ES as distributional scoring rules can be found in Appendix A.
[2] For a more detailed explanation of the copula structure, see Appendix C

one-time cost. The time for creating a single, new period is around 30 minutes. This makes the project feasible for common production workflows.

5 Results

To maintain a reasonably stable predictor set for the estimation, the out-of-sample period stretches from February 2013 to June 2025, with quarterly YoY forecasts made every month. For all results, it should be noted that due to the immense size of the shock, the early COVID period is excluded from the results as it would heavily skew results and scores. This leaves me with 132 out-of-sample observations. For evaluation, I use the third release of the variables, usually available 90 days after the quarter ends, to keep a policy-relevant timeframe for assessment.

For evaluation, I consider both univariate and multivariate scoring rules, namely the continuous ranked probability score (CRPS) and its multivariate counterpart, the energy score (ES) next to the multivariate variogram score (VS)[1]. For CRPS and ES, lower is better.

5.1 Benchmarks

To have baselines to compare the HiDVAR to, a Bayesian VAR (BVAR) based on the prior selection by Giannone et al. (2015) with stochastic volatility (BVAR-SV) is used, which also uses the same factor information set as the main model. Thus, you may also call this a FAVAR with a GLP-style prior. Furthermore, a distributional regression in the sense of a Penalized Quantile Regression with a student-t copula (PQR) is used. As a contemporary machine learning benchmark, I use the Quantile Random Forest Model with Diffusion Indices (QRFARDI) popularized by Goulet Coulombe et al. (2022). Finally, I directly compare it to the results of the FRBNY model, which I extracted from their website. All four benchmarks are described in Appendix D.

5.2 Calibration

Before taking a deep dive into the out-of-sample scores, one can look into how the model performs in calibration exercises by comparing nominal and empirical coverage. This can be done by calculating how many of the empirical observations are captured by which density bands of the forecasts. An optimally calibrated model would run along the 45°-line. Figure 4 indicates that inflation and GDP are fairly well calibrated considering the small out-of-sample size except for the upper ends of the distribution. Unemployment, on the other hand,

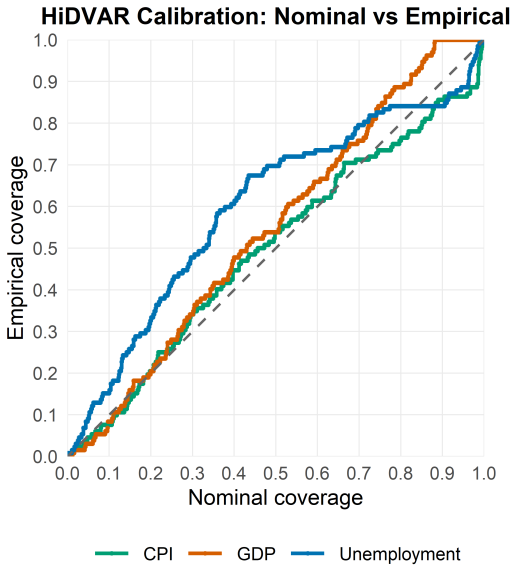


Figure 4: Nominal vs Empirical Coverage for HiDVAR

[1] Exact definitions of the scoring rules can be found in Appendix A.

is slightly underconfident, meaning that
the distribution is too conservative but
is marginally too confident on CPI and GDP at the upper end. Overall, this is a satisfying, but far from perfect result. For further calibration diagnostics, I also report PIT histograms and Kolmogorov-Smirnov (Kolmogorov, 1933; Smirnov, 1948) tests for all models. Probability

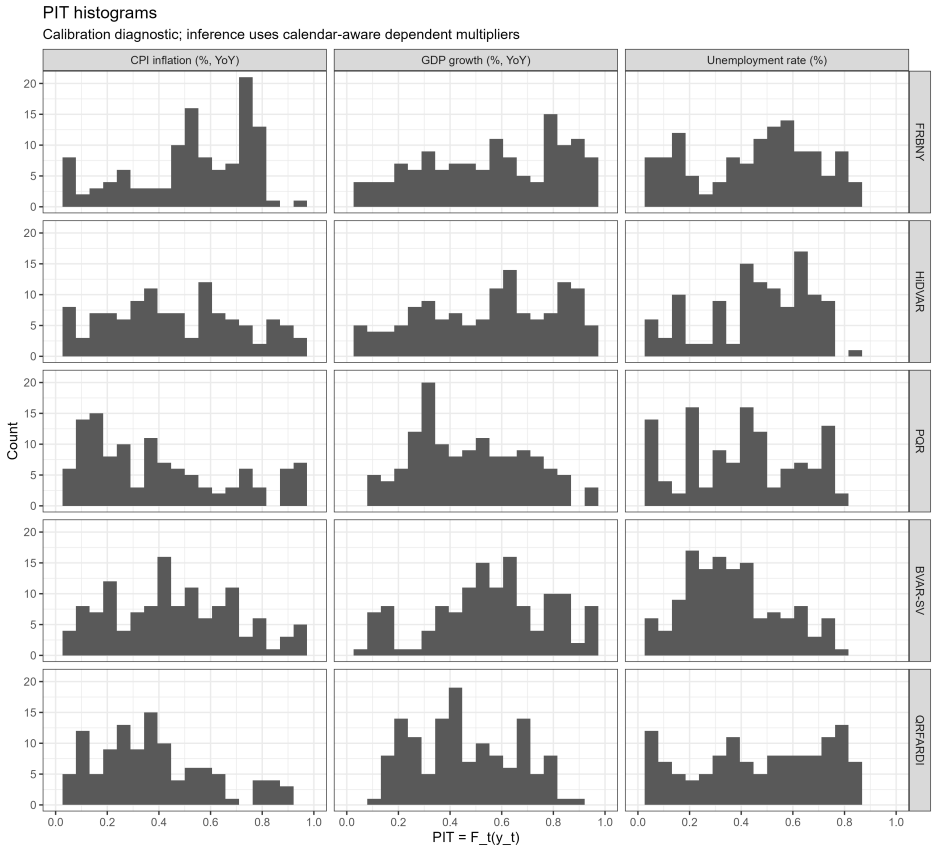


Figure 5: PIT Histograms for HiDVAR and FRBNY forecasts (N=132)

integral transforms (PITs), used to assess probabilistic calibration, are defined as

$$u_t = F_t(y_t), \tag{4}$$

and should be uniformly distributed on $[0, 1]$ under correctly specified continuous predictive distributions. Figure 5 presents PIT histograms, while Table 1 reports Kolmogorov–Smirnov statistics and dependence-adjusted bootstrap p-values. The bootstrap accounts for dependence across overlapping forecasts and clusters forecasts sharing a target quarter. The tests assess marginal PIT uniformity; non-rejection does not establish full conditional calibration.

For CPI inflation, HiDVAR exhibits the smallest deviation from uniformity among the five models, with a KS statistic of 0.108 ($p = 0.595$), followed by BVAR-SV at 0.125 ($p = 0.468$). Although FRBNY has a larger statistic of 0.216, uniformity is not rejected at conventional significance levels ($p = 0.169$). In contrast, uniformity is rejected at the 10% level for PQR ($D = 0.238$, $p = 0.083$) and at the 5% level for QRFARDI ($D = 0.255$, $p = 0.040$). These diagnostics are consistent with comparatively well-calibrated marginal inflation forecasts from HiDVAR, although the individual uniformity tests do not establish statistically significant calibration differences between models.

For GDP growth, uniformity is not rejected at the 10% level for any model. HiDVAR and FRBNY have the smallest KS statistics, at 0.113 and 0.120, respectively, followed by PQR (0.144), QRFARDI (0.188), and BVAR-SV (0.192).

For unemployment, QRFARDI and FRBNY exhibit the smallest deviations from uniformity, with statistics of 0.146 ($p = 0.361$) and 0.168 ($p = 0.262$), respectively. Uniformity is also not rejected for PQR ($p = 0.101$), whereas it is rejected at the 10% level for HiDVAR ($p = 0.063$) and at the 1% level for BVAR-SV ($p = 0.001$). Thus, HiDVAR's favorable inflation diagnostics do not extend uniformly across outcomes.

Table 1: Kolmogorov–Smirnov Tests (Model forecasts, $N = 132$)

FRBNY	HiDVAR	PQR	BVAR-SV	QRFARDI
CPI inflation (%, YoY)				
0.216	0.108	0.238	0.125	0.255
(p=0.169)	(p=0.595)	(p=0.083)	(p=0.468)	(p=0.040)
GDP growth (%, YoY)				
0.120	0.113	0.144	0.192	0.188
(p=0.599)	(p=0.511)	(p=0.286)	(p=0.134)	(p=0.128)
Unemployment rate (%)				
0.168	0.238	0.229	0.283	0.146
(p=0.262)	(p=0.063)	(p=0.101)	(p=0.001)	(p=0.361)

To complement the graphical calibration analysis, Table 2 reports the energy-score-based multivariate calibration tests of Knüppel et al. (2022). The generalized Box transform test does not reject calibration at the 10% level for any of the five models. The corresponding p-values range from 0.204 for BVAR-SV to 0.762 for PQR, with HiDVAR, FRBNY, and QRFARDI yielding p-values of 0.435, 0.271, and 0.486, respectively.

The complementary entropy test likewise provides no statistically significant evidence of overconfidence or underconfidence for any model. HiDVAR and FRBNY have entropy statistics close to zero, at -0.173 ($p = 0.863$) and 0.318 ($p = 0.751$), respectively. The statistics for BVAR-SV, PQR, and QRFARDI are negative, but none is significant at the 10% level. Their signs therefore do not provide sufficient evidence to characterize these forecasts as underconfident. Taken together, these results reveal no significant departures from the calibration restrictions examined by either test across the five models.

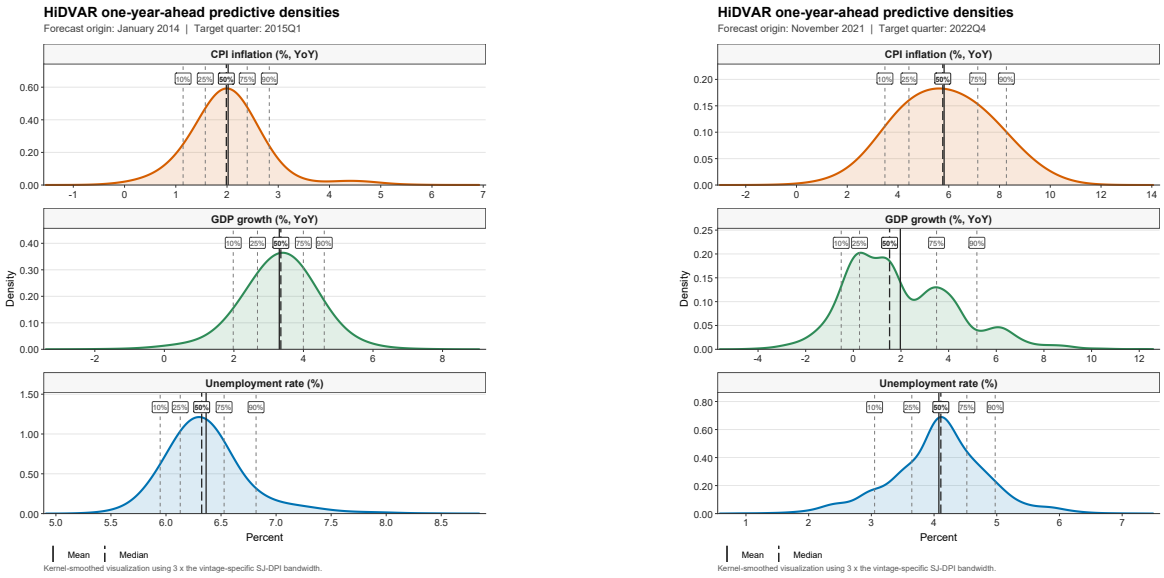
Table 2: Energy-Score-based multivariate calibration tests

Model	GBT statistic	p-value	Entropy t	p-value	Entropy diagnosis
HiDVAR	3.79	0.435	−0.17	0.863	No significant misspecification
FRBNY	5.16	0.271	0.32	0.751	No significant misspecification
BVAR-SV	5.93	0.204	−1.57	0.115	No significant misspecification
PQR	1.86	0.762	−0.47	0.641	No significant misspecification
QRFARDI	3.44	0.486	−1.29	0.199	No significant misspecification

Notes: $N = 132$. Tests use the standardized energy score and heteroskedasticity- and autocorrelation-consistent (HAC) covariance estimates with 11 lags. The generalized Box transform test assesses whether the first four moments of the energy-score-based probability integral transform equal the corresponding moments of a uniform distribution. The entropy test examines whether the mean difference between realized and forecast-implied score uncertainty equals zero. A significantly positive entropy statistic indicates overconfidence, whereas a significantly negative statistic indicates underconfidence. No significant misspecification denotes non-rejection by the entropy test at the 10% significance level.

5.3 Distributional Output

At first, I analyze how much the distributions actually change under the HiDVAR over time. Figure 6 illustrates the predictive densities for selected forecast vintages. The predictive distributions vary substantially over time, especially for inflation. It can be observed that the



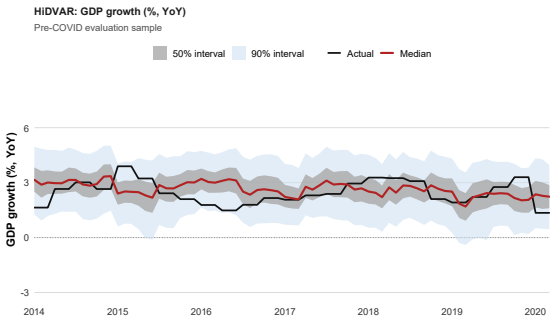
(a) January 2014 **(b)** November 2021

Figure 6: 1-year ahead densities by the HiDVAR for different vintages

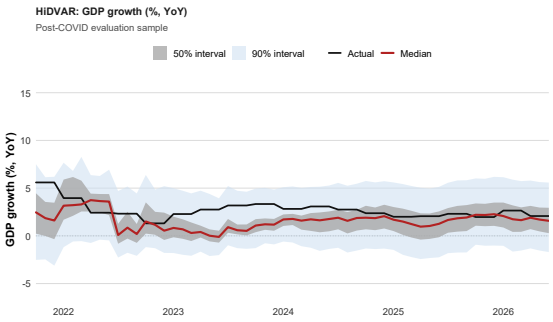
shape of the distributions changes drastically throughout the out-of-sample period. Figure 6 illustrates fairly regular distributions in January 2014. However, not only can the distributions change greatly over time in support and size, they can also change shape substantially. The most extreme shape changes occurred during the COVID period, where extreme signals exemplify the uncertainty around the shape of the forecasts. Furthermore, multimodality can be found for GDP in November 2021, where there are several peaks with a skew towards low growth. In general, this exemplifies that the distributions of the HiDVAR react to macroeconomic uncertainty and change size, shape, and location with respect to that.

5.4 Out-of-sample Evaluation

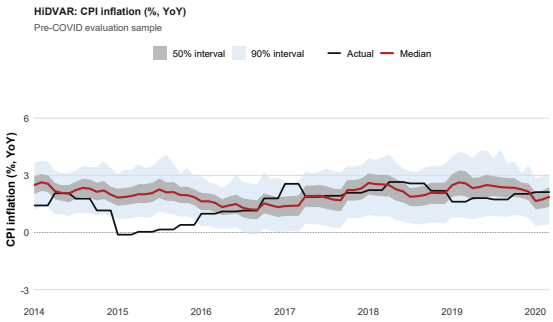
The main results are similar to the findings about calibration. Mainly, the HiDVAR does an excellent job in predicting CPI inflation, compared to both the FRBNY model as well as standard alternatives, while being very similar in performance to the FRBNY when it comes to GDP and unemployment. Concretely, it can reduce CPI CRPS by up to 27.7%, ES by 21.8% and VS by 25.1%. In general, the errors for unemployment are the lowest, followed by GDP and inflation due to the low volatility of the unemployment rate and the (mostly) unexpected spike in prices.



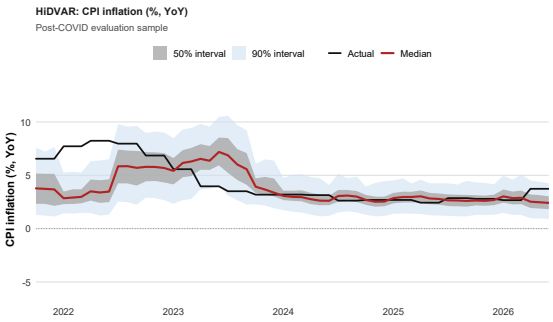
(a) HiDVAR pre-COVID GDP



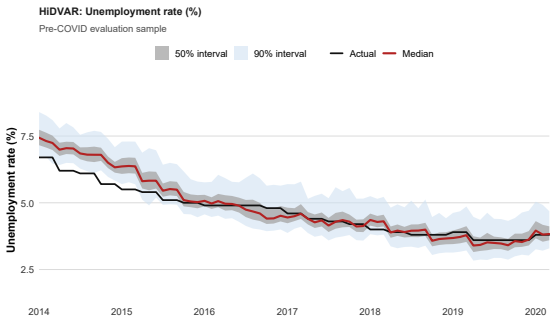
(b) HiDVAR post-COVID GDP



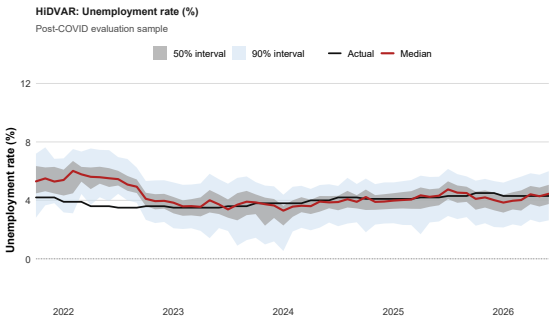
(c) HiDVAR pre-COVID CPI



(d) HiDVAR post-COVID CPI



(e) HiDVAR pre-COVID Unemployment



(f) HiDVAR post-COVID Unemployment

Figure 7: HiDVAR 1-year ahead densities pre- and post-COVID against actual, forecasts shifted to match Target

The results further indicate that HiDVAR's performance cannot be attributed solely to its information set, because the principal model-based benchmarks use a comparable factor set. For GDP, the PQR and QRFARDI perform best. This aligns with the key characteristics of the HiDVAR. As can be inferred from Figure 3a, GDP is relatively symmetric and normally distributed, making the advantages of the nonparametric method rather small. For

Table 3: Probabilistic Evaluation Scores, $N = 132$

	FRBNY	HiDVAR	PQR	BVAR-SV	QRFARDI
GDP CRPS	0.717	0.670	0.565	0.875	**0.552**
CPI CRPS	1.008	**0.729**	0.972	0.920	0.909
Unemployment CRPS	**0.293**	0.315	0.361	0.413	0.311
Energy Score	1.040	**0.813**	0.968	1.050	0.911
Variogram Score	1.298	**0.972**	1.169	1.489	1.175

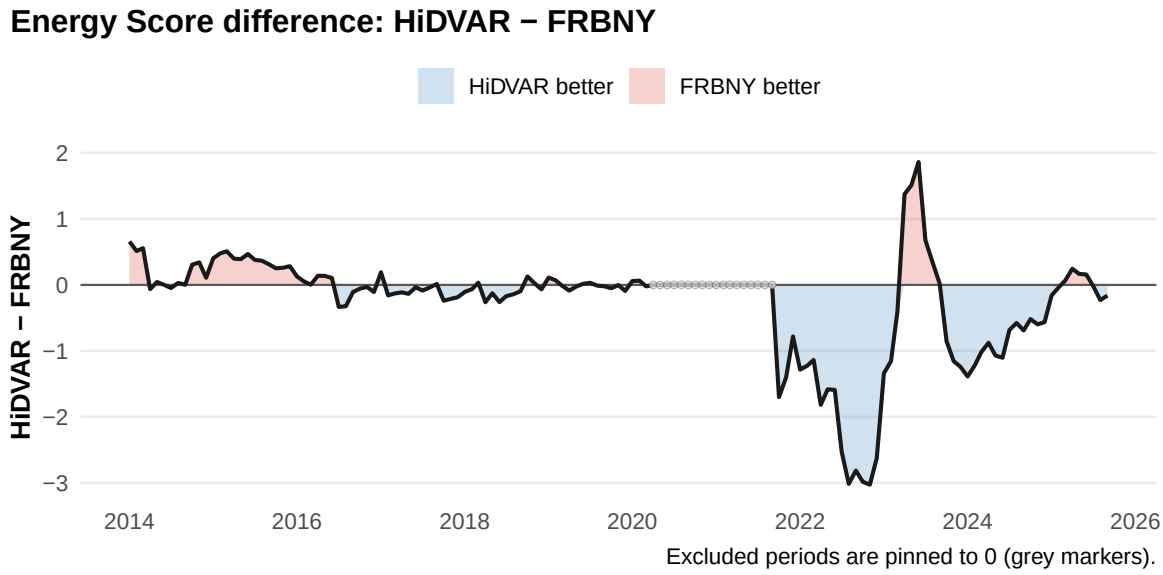
Notes: Lower values indicate better predictive performance. Bold entries denote the best-performing model in each row. The COVID evaluation window is excluded.

unemployment, FRBNY, HiDVAR, and QRFARDI show quite similar performance. For the multivariate scores, the HiDVAR dominates. It provides much lower errors for both energy score and variogram score than all benchmarks, highlighting the model's main objective, multivariate predictions. To take a closer look at where gains and losses occur, Figure 8 reports time-varying score differentials computed as HiDVAR - FRBNY. Negative values (blue shading) indicate that HiDVAR attains lower (better) loss, while positive values (red shading) favor FRBNY. The upper panel uses the multivariate energy score, summarizing relative performance for the joint predictive distribution, whereas the lower panels decompose the comparison into univariate CRPS differentials for CPI inflation, GDP growth, and unemployment. Differences are typically small in tranquil periods, suggesting broadly comparable accuracy. The most pronounced improvements for HiDVAR occur during the post-pandemic inflation episode, where the CPI CRPS differential turns strongly negative and drives the large negative energy score gap, while GDP and unemployment show smaller, more mixed contributions with brief intervals in which FRBNY performs better.

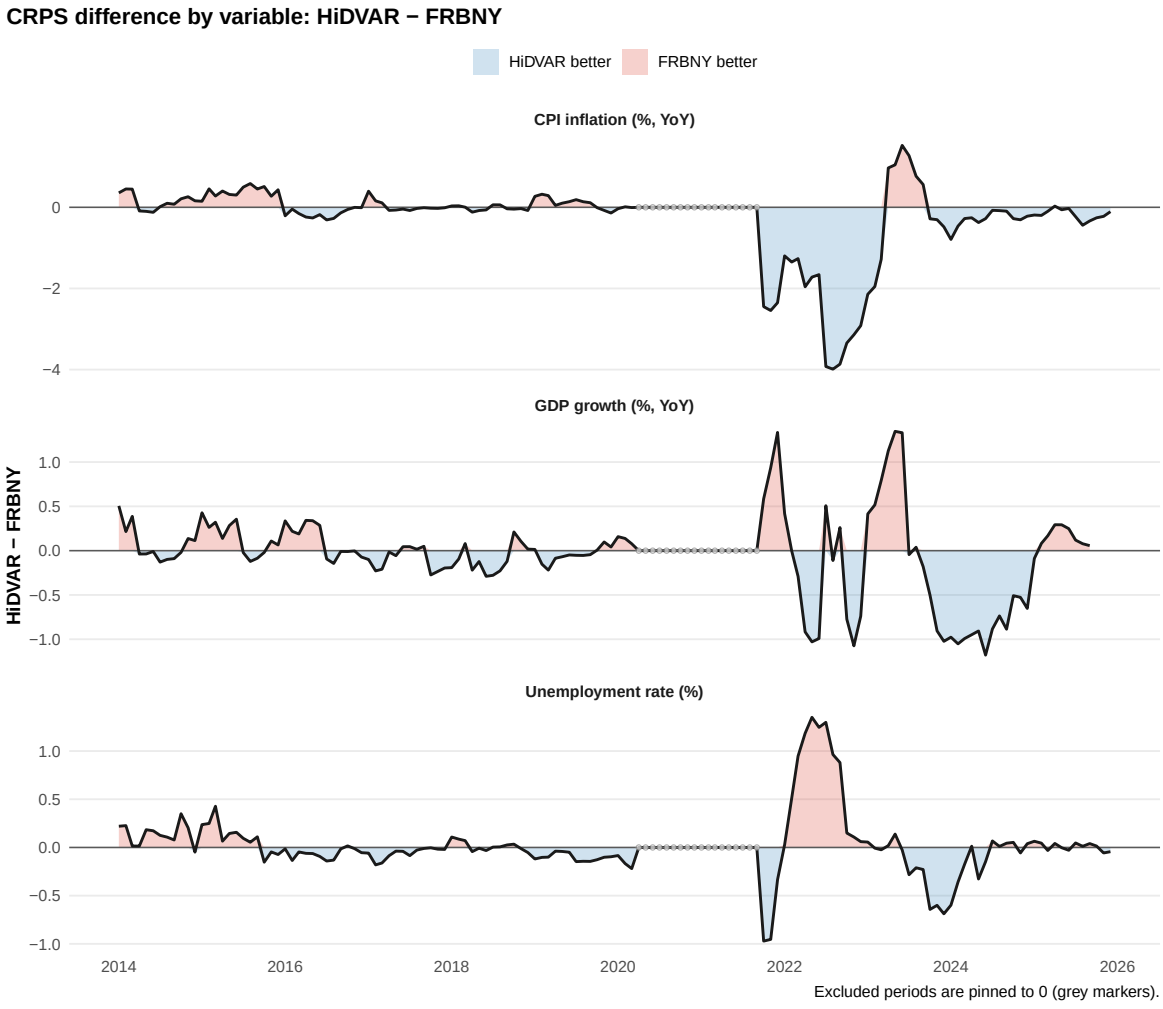
To better understand differences in CPI forecasting performance, I decompose the CRPS into components capturing miscalibration, discrimination, and intrinsic uncertainty. This can help alleviate the issue of static forecasts with wide distributions performing better in volatile regimes while offering little to no discrimination or sharpness. Following Gneiting and Resin (2023) and Arnold et al. (2024), the average score over forecast–outcome pairs (F_t, y_t), $t = 1, \ldots, T$, can be written as

$$\overline{\text{CRPS}} = \frac{1}{T} \sum_{t=1}^{T} \text{CRPS}(F_t, y_t) = \text{MCB} - \text{DSC} + \text{UNC}, \tag{5}$$

where MCB denotes the miscalibration component, DSC the discrimination component, and UNC the uncertainty component. Intuitively, MCB measures the extent to which issued predictive distributions fail to be conditionally calibrated, DSC captures the ability of the forecasts to distinguish between different future states of the economy, and UNC reflects the intrinsic variability of the realized outcome process. Lower values of MCB and higher values of DSC are desirable, whereas UNC is a property of the forecast target rather than of the forecasting model. The application to all models for inflation, where the main gains are coming from, can be found in Table 4. The decomposition makes it clearer where the HiDVAR outperforms the benchmarks. The decomposition suggests that HiDVAR's CPI gains do not primarily reflect lower miscalibration, but substantially stronger discrimination between low- and high-inflation-risk states. The result is strong given that sharpness and discrimination are especially relevant for policy decisions, something which the FRBNY



(a) Energy score difference (HiDVAR – FRBNY).



(b) CRPS differences by variable (HiDVAR – FRBNY).

Figure 8: Score differentials between HiDVAR and FRBNY over time. Negative values indicate lower (better) scores for HiDVAR. Excluded periods are pinned to zero (grey markers).

Table 4: Inflation CRPS decomposition by model, $N = 132$

Model	$\overline{\text{CRPS}}$	MCB	DSC	UNC
FRBNY	1.008	0.221	0.226	
HiDVAR	**0.729**	0.213	**0.496**	
PQR	0.972	**0.186**	0.225	1.01
BVAR-SV	0.920	0.229	0.319	
QRFARDI	0.909	0.327	0.429	

Notes: The mean CRPS is decomposed as $\overline{\text{CRPS}} = \text{MCB} - \text{DSC} + \text{UNC}$. The uncertainty component is identical across models because they are evaluated on the same realizations. Bold entries indicate the lowest mean CRPS and miscalibration (MCB), and the highest discrimination (DSC). Rounding may cause small discrepancies in the displayed decomposition.

model is especially lacking for inflation. This also supports the strength of the model in choosing its own distribution each period.

To further address where in the distribution gains occur, I analyze the Interval Score, introduced by Gneiting and Raftery (2007) for the 50% and 90% intervals only. The rule is explained in detail in Appendix A.

Table 5: Interval Scores, $N = 132$

	IS_{50}			IS_{90}		
Model	GDP	CPI	UR	GDP	CPI	UR
FRBNY	3.24	4.32	**1.33**	5.46	12.20	2.95
HiDVAR	3.08	**3.28**	1.46	**5.11**	**6.73**	3.21
PQR	2.48	4.39	1.59	5.92	8.90	3.89
BVAR-SV	4.07	4.16	1.81	6.72	7.41	3.42
QRFARDI	**2.43**	4.06	1.40	6.31	8.24	**2.65**

Prediction intervals provide an interpretable summary of predictive uncertainty. Table 5 reports interval scores for the 50% and 90% prediction intervals, with lower values indicating a better balance between interval width and penalties for observations falling outside the interval. For GDP, QRFARDI achieves the lowest score at the 50% level, while HiDVAR performs best at the 90% level. For unemployment, FRBNY and QRFARDI lead at the respective coverage levels. HiDVAR's advantage is most consistent for inflation, where it achieves the lowest scores at both levels. Relative to FRBNY, HiDVAR reduces the inflation interval score by approximately 24% for the 50% interval and 45% for the 90% interval. The larger relative improvement at the higher coverage level is consistent with better characterization of inflation uncertainty beyond the center of the distribution, although interval scores alone do not establish superior forecasting of tail events.

As the main crisis during the out-of-sample period was unforeseen inflation, I also look at logistic tail-weighted CRPS[2]. Here, the differences become even more striking. The FRBNY model now even loses out against the PQR benchmark. The HiDVAR lowers errors compared

[2]For more information, see Appendix A.

to the FRBNY model by 31% for TWCRPS. The gains over the model benchmarks are also substantial but smaller in size.

Table 6: Logistic Tail-Weighted CRPS, $N = 132$

	FRBNY	HiDVAR	PQR	BVAR-SV	QRFARDI
Logistic TWCRPS	0.580	**0.403**	0.549	0.485	0.538

To assess whether differences in predictive accuracy are statistically significant, I employ one-sided Diebold and Mariano (1995) (DM) tests based on CRPS, energy score, and variogram score differentials. Appendix A provides the detailed test specification. Score differentials are defined as HiDVAR minus the respective benchmark, so negative DM statistics favor HiDVAR. Inference accounts for serial dependence using an eleven-lag Bartlett HAC estimator.

Table 7 provides modest evidence of improved predictive accuracy relative to FRBNY. HiDVAR achieves lower energy and variogram scores, with both differences significant at the 10% level ($p = 0.075$ and $p = 0.069$, respectively). Among the univariate comparisons, the inflation CRPS improvement is also significant at the 10% level ($p = 0.099$), whereas the tests do not establish improvements for GDP growth or unemployment.

The evidence against the remaining benchmarks varies across variables and scores. HiDVAR outperforms BVAR-SV across all reported criteria at the 1% level. Relative to PQR, the improvements are significant for inflation, unemployment, and both multivariate scores, but not for GDP growth. Against QRFARDI, HiDVAR achieves a lower inflation CRPS at the 5% significance level and lower energy and variogram scores at the 10% level. QRFARDI has the lower average GDP CRPS, while the unemployment comparison provides no evidence of an advantage for HiDVAR. Overall, the results indicate that HiDVAR's most consistent gains concern inflation and joint predictive accuracy, with stronger statistical evidence against BVAR-SV and PQR than against FRBNY and QRFARDI.

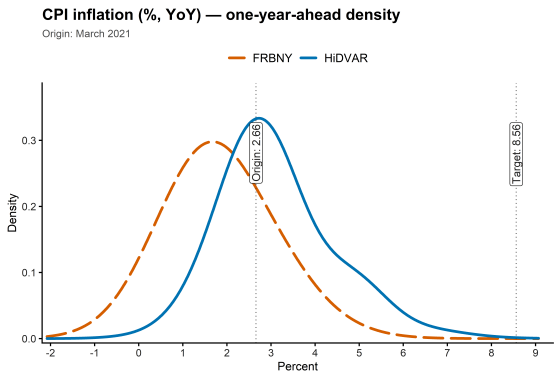
Table 7: Diebold–Mariano Tests for Score Differentials (HiDVAR Minus Benchmark), $N = 132$

	FRBNY		BVAR-SV		PQR		QRFARDI	
	DM	p-value	DM	p-value	DM	p-value	DM	p-value
Panel A: Univariate CRPS								
GDP growth	−0.588	0.278	−2.97	0.0015***	1.38	0.916	1.89	0.971
CPI inflation	−1.28	0.099*	−2.70	0.003***	−3.55	< 0.001***	−1.82	0.034**
Unemployment	0.368	0.644	−3.76	< 0.001***	−1.39	0.082*	0.218	0.586
Panel B: Multivariate Scores								
Energy Score	−1.44	0.075*	−3.66	< 0.001***	−2.49	0.006***	−1.55	0.061*
Variogram Score	−1.49	0.069*	−2.71	0.003***	−2.05	0.020**	−1.35	0.088*

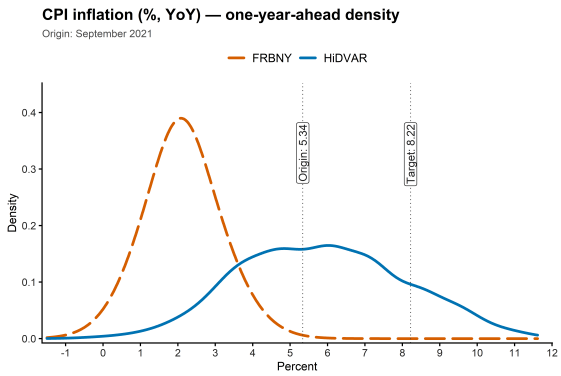
Notes: Score differentials are defined as HiDVAR minus the respective benchmark, so negative statistics favor HiDVAR. All reported p-values are one-sided, testing the alternative that the expected score differential is negative (HiDVAR outperforms the benchmark). Inference uses an eleven-lag Bartlett HAC covariance estimator. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$. The COVID evaluation window is excluded.

To give a visual example of how and where the inflation score differences occur, I directly compare the predictive densities in Figure 9. In September 2021, the difference becomes

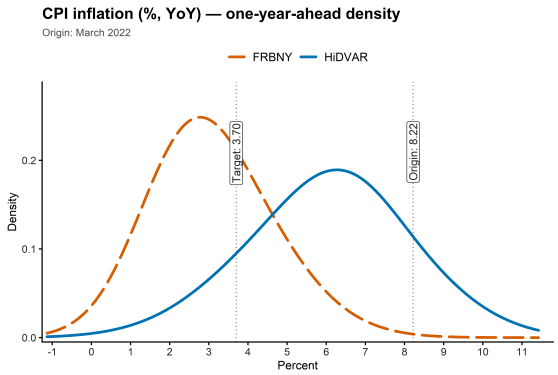
most apparent. Added are the inflation at the time when the forecast was made (~5.34%), and the targeted realized inflation one year later (~8.22%). The FRBNY model is predicting a return to low inflation way too early. Even more concretely, the FRBNY model set the current inflation value into the 99.85th percentile of its distribution. This pattern is consistent with strong mean-reversion assumptions and insufficient upper-tail mass. In comparison, the HiDVAR placed it at the 42.5th percentile, setting on even higher inflation in the future, which proved to be correct. The observed value is even more extreme. Although even the HiDVAR deemed it rather unlikely at the 86th percentile, FRBNY-OaR sees a virtually 0% chance of inflation being that high. A crucial miscalculation given that inflation peaked at 8.60% in 2022Q2. The HiDVAR also put much more weight on high inflation in March 2021, whereas the FRBNY model put around as much weight on deflation as on inflation surpassing 3.5%. The picture evens out more when looking at the disinflation period. In March 2022, when inflation peaked at around 8.22%, the FRBNY model crucially still again provides a strong mean reversion back into the realms of 2%, only that now, inflation actually fell. However, it still predicts a too fast recovery, putting the later observed inflation of 3.70% in the 95th percentile. In general, the FRBNY model forecasts only changes mildly, reflecting the mean reversion that many institutions predicted (Figure 1). The HiDVAR, on the other hand, provides many different shapes for inflation forecasts, adapting to real-time conditions.



(a) Origin: March 2021



(b) Origin: September 2021



(c) Origin: March 2022

Figure 9: 1-year ahead densities for CPI inflation of HiDVAR and FRBNY

5.5 Ablation of Model Components

The HiDVAR combines two extensions over Wang et al. (2026): a high-dimensional factor information set and an extended multivariate dependence structure based on recursive contemporaneous conditioning and copula-based simulation. To assess whether these components materially contribute to predictive performance, I consider two restricted versions of the model.

First, I remove the common factors from the forecasting equations while retaining the distribution-regression, regularization, triangular-factorization, and copula structure. This specification conditions only on the outcome variables and their lags and therefore isolates the contribution of the high-dimensional information set.

Second, I retain the baseline information set but remove the multivariate forecasting structure. GDP growth, CPI inflation, and unemployment are then estimated as separate distributional regressions without recursive contemporaneous conditioning, while forecast draws are generated independently rather than through the copula. This specification therefore assesses the combined contribution of the triangular factorization and the dependence structure used for joint path simulation.

Table 8: Ablation of HiDVAR model components

Specification	GDP CRPS	CPI CRPS	UR CRPS	ES	VS
Baseline HiDVAR	0.670	**0.729**	**0.315**	**0.813**	**0.972**
Without high-dimensional factors	0.642	0.855	0.431	0.937	1.155
Without multivariate structure	0.616	0.872	0.465	0.946	1.154
with skew-t parametric smoothing	**0.580**	0.943	0.320	0.936	1.100

Notes: The factor ablation removes the common factors while retaining the remaining HiDVAR architecture. The multivariate ablation retains the baseline information set but removes both recursive contemporaneous conditioning and copula-based dependence, such that the three outcome distributions are modeled and simulated independently. The last ablation differs from the baseline by fitting 19 quantiles from HiDVAR into a skew-t distribution instead of using the nonparametric forecasts. $N = 132$.

Table 8 reveals a strikingly consistent pattern across the two exercises. Neither the high-dimensional information set nor the multivariate architecture improves marginal forecasting performance uniformly. For GDP growth, removing the factors lowers CRPS from 0.670 to 0.642, while the fully univariate specification lowers it further to 0.616. Also, smoothing the forecasts by putting them into a skew-t distribution, similar to the approach by Adams et al. (2021) decreases it even more to 0.580. Hence, the additional information, nonparametric density, and dependence structure introduce some estimation complexity that does not translate into improved GDP density forecasts.

The opposite result emerges for inflation and unemployment. Removing the factor information increases CPI CRPS from 0.729 to 0.855 and unemployment CRPS from 0.315 to 0.431. Removing the multivariate structure produces a similar, somewhat larger deterioration, with CRPS values of 0.872 and 0.465, respectively. Relative to the baseline specification, this corresponds to approximately 17–19% higher CPI CRPS and 37–48% higher unemployment CRPS. For the skew-t version, inflation errors increase to 0.943, by 29%, showing the negative effect parametric distribution fitting can have on inflation forecasts.

The same pattern is visible in the joint forecast. The energy score rises from 0.813 under the baseline HiDVAR to 0.937 without the factor information, 0.946 without the multivariate structure, and 0.936 with skew-t smoothing, corresponding to deteriorations of approximately 15–16%. Similar sized results hold for the variogram score. The gains of the full model therefore cannot be attributed to a single modeling component. Both the broader macroeconomic information set and the explicit modeling of dependence contribute to the accuracy of the joint predictive distribution.

At the same time, the GDP results caution against interpreting either extension as mechanically beneficial. Their value is concentrated in the variables and forecasting objects most relevant for the paper's risk-monitoring application: inflation, labor-market conditions, and their joint evolution. This finding is consistent with the main out-of-sample results, where HiDVAR's strongest advantage is likewise concentrated in inflation-risk assessment rather than in uniformly superior forecasts across all variables.

5.6 Bandwidth Choice and Sensitivity

The HiDVAR produces predictive distributions in the form of simulated forecast draws. For graphical presentation, these draws are converted into continuous density estimates using kernel density estimation. The bandwidth then determines the degree of smoothing applied in this step. A small bandwidth closely follows the empirical distribution of the simulated draws and can therefore reveal features such as local asymmetry or heavy multimodality shoulders. However, it may also translate finite-simulation noise or small clusters of draws into economically uninformative local peaks. A larger bandwidth produces a smoother density by averaging over a wider neighborhood, improving visual clarity at the cost of potentially attenuating genuine distributional features.

This trade-off is particularly relevant for policy applications. While the ability to represent non-Gaussian and potentially multimodal forecast distributions is an important advantage of the HiDVAR, policy communication requires densities that remain interpretable and do not suggest distinct economic regimes merely because of minor irregularities in the simulated draws. Also, for policy communication, multimodality can create issues in readability especially for non-technical readers, so providing a smoothed version is essential for production routines. I therefore use three times the vintage-specific Sheather–Jones direct plug-in bandwidth for the graphical presentation of the predictive densities (Sheather & Jones, 1991). Importantly, this smoothing is applied to the simulated forecast draws themselves rather than only to the plotted density curve. The bandwidth choice can consequently affect both the visual appearance of the forecasts and, in principle, their evaluated predictive scores.

Figures 10–13 assess this trade-off systematically. Figure 10 reports the percentage change in the mean CRPS relative to the unsmoothed empirical-draw forecast. At the preferred bandwidth multiplier of three, the effect is negligible for CPI inflation and remains limited for the other variables. The CPI CRPS is virtually unchanged, while the increases for GDP growth and unemployment are approximately 1.2% and 0.6%, respectively. More aggressive smoothing eventually leads to larger deteriorations, particularly for GDP, but these occur primarily at bandwidth multipliers beyond the preferred value.

The same conclusion holds for the joint predictive distribution. As shown in Figure 11, the mean energy score increases by only approximately 0.7% at the preferred bandwidth. The deterioration remains modest relative to the unsmoothed joint distribution, while substantially

stronger smoothing at multipliers of four or six produces more noticeable losses. Thus, the selected bandwidth lies in a region in which the evaluated probabilistic accuracy remains close to that of the original empirical-draw forecasts.

At the same time, smoothing has a pronounced effect on apparent multimodality. Figure 13 reports the share of forecast densities classified as multimodal after ignoring local peaks below 5% of the global density maximum. With only limited smoothing, nearly all vintages are classified as multimodal, suggesting that many of the detected peaks reflect small-scale irregularities rather than economically distinct forecast regimes. At the preferred multiplier of three, the multimodal share falls to approximately 10% for CPI inflation and slightly above 20% for GDP growth and unemployment. Further smoothing reduces these shares only gradually, while imposing increasingly visible score losses.

Taken together, the results suggest that the preferred bandwidth provides a useful compromise. It removes a large fraction of visually distracting and potentially spurious local modes, while leaving both the marginal CRPS and the multivariate energy score largely unchanged. The reported density shapes should therefore be interpreted as smoothed representations of the underlying forecast draws rather than as additional parametric restrictions imposed during model estimation.

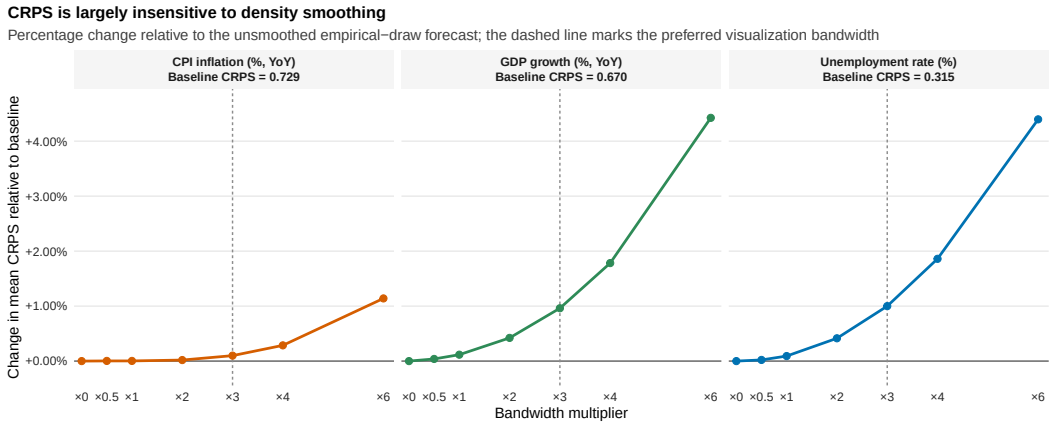


Figure 10: Sensitivity of marginal CRPS to the density bandwidth. Percentage changes are measured relative to the unsmoothed empirical-draw forecasts. The dashed vertical line denotes the chosen bandwidth multiplier of three.

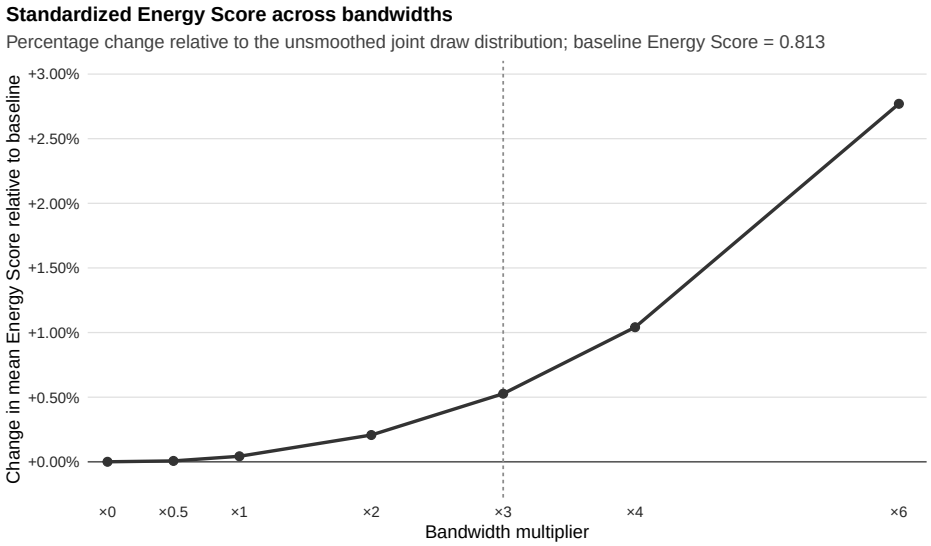


Figure 11: Sensitivity of the multivariate energy score to the density bandwidth. Percentage changes are measured relative to the unsmoothed joint draw distribution. The dashed vertical line denotes the chosen bandwidth multiplier of three.

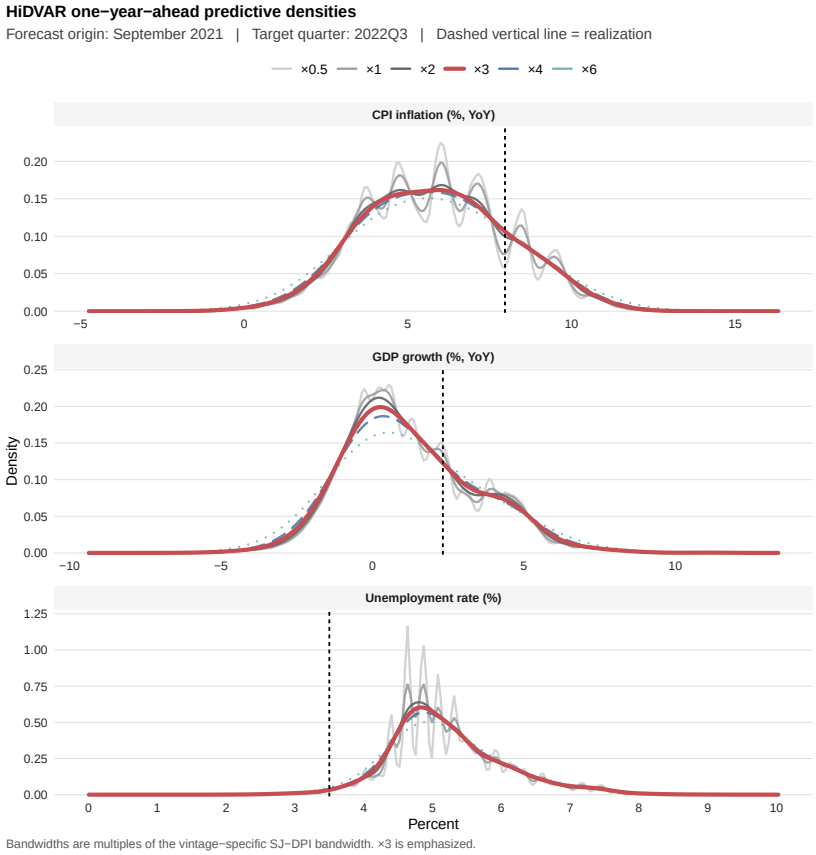


Figure 12: Bandwidth Sensitivity, exemplary for September 2021

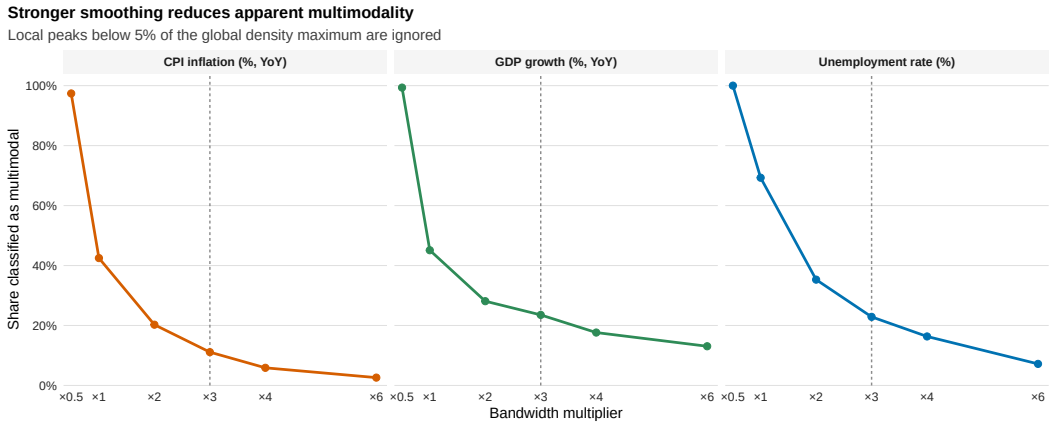


Figure 13: Density Bandwidth and Apparent Multimodality

Notes: A forecast density is classified as multimodal if it contains more than one local peak exceeding 5% of the global density maximum. The dashed vertical line marks the selected bandwidth multiplier of three.

6 Risk Monitoring and Policy Translation

6.1 Macro Risk Dashboard

Other than the standard forecasting results, the structure also allows for creating other probabilistic measures which are not possible using just point forecasts or univariate densities.

6.1.1 Macroeconomic Scenario Probabilities

A useful advantage of the HiDVAR framework is that it does not only provide marginal forecast densities, but full simulated paths for the macroeconomic variables in one unified system. Inspired by the recommendations in the BoE review report by Bernanke (2024), I use these paths to construct one-year-ahead probabilities of policy-relevant macroeconomic scenarios. Instead of evaluating recession or stagflation risk in isolation, each simulated path is assigned to exactly one mutually exclusive macro-risk regime. This makes the resulting probabilities directly interpretable as scenario shares.

For each forecast origin t, let $i = 1, \dots, N$ index the simulated paths, with $N = 10{,}000$. The classification is based on four-quarter-ahead GDP growth, CPI inflation, and unemployment dynamics. I define a technical recession indicator as

$$R_{i,t} = \mathbb{1}\left\{\exists h \in \{1,2,3\} : g^{qoq}_{i,t+h} < 0 \text{ and } g^{qoq}_{i,t+h+1} < 0\right\},$$

that is, a simulated path is classified as containing a recession if it contains two consecutive quarters of negative GDP growth within the one-year forecast window. Inflation pressure is defined as

$$I_{i,t} = \mathbb{1}\left\{\sum_{h=1}^{4} \mathbb{1}\left(\pi^{yoy}_{i,t+h} > 3\right) \geq 2\right\},$$

so that inflation must exceed 3% in at least two of the next four quarters. This avoids classifying short-lived price spikes as a persistent inflation regime. Weak growth is defined by

$$W_{i,t} = \mathbb{1}\left\{g^{yoy}_{i,t+4} < 0.75\right\},$$

and labor-market stress is defined as

$$U_{i,t} = \mathbb{1}\left\{\max_{h=1,\dots,4}(u_{i,t+h} - u_t) \geq 0.5\right\}.$$

Hence, a simulated path exhibits real-side stress if it contains either a technical recession, a material increase in unemployment, or weak terminal year-on-year GDP growth.

The thresholds are chosen to have an empirical and economic interpretation. The weak-growth threshold of 0.75% is close to the lower tail of the historical US growth distribution: since 1990, the 10th percentile of year-on-year GDP growth is approximately 0.64%.

The six scenarios are now defined as follows:

Stable expansion:	$I_{i,t} = 0,\ R_{i,t} = 0,\ U_{i,t} = 0,\ W_{i,t} = 0,\ \pi^{yoy}_{i,t+4} \leq 3,\ g^{yoy}_{i,t+4} \geq 0.75,$
Inflationary expansion:	$I_{i,t} = 1,\ R_{i,t} = 0,\ U_{i,t} = 0,\ W_{i,t} = 0,$
Stagflation risk:	$I_{i,t} = 1$ and $(R_{i,t} = 1$ or $U_{i,t} = 1$ or $W_{i,t} = 1),$
Recessionary disinflation:	$I_{i,t} = 0$ and $(R_{i,t} = 1$ or $U_{i,t} = 1),$
Weak-growth disinflation:	$I_{i,t} = 0,\ R_{i,t} = 0,\ U_{i,t} = 0,\ W_{i,t} = 1,$
Mixed / transition:	all remaining paths.

The probability of scenario s at origin t is then simply the share of simulated paths assigned to that scenario,

$$\widehat{p}_{s,t} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\{C_{i,t} = s\}.$$

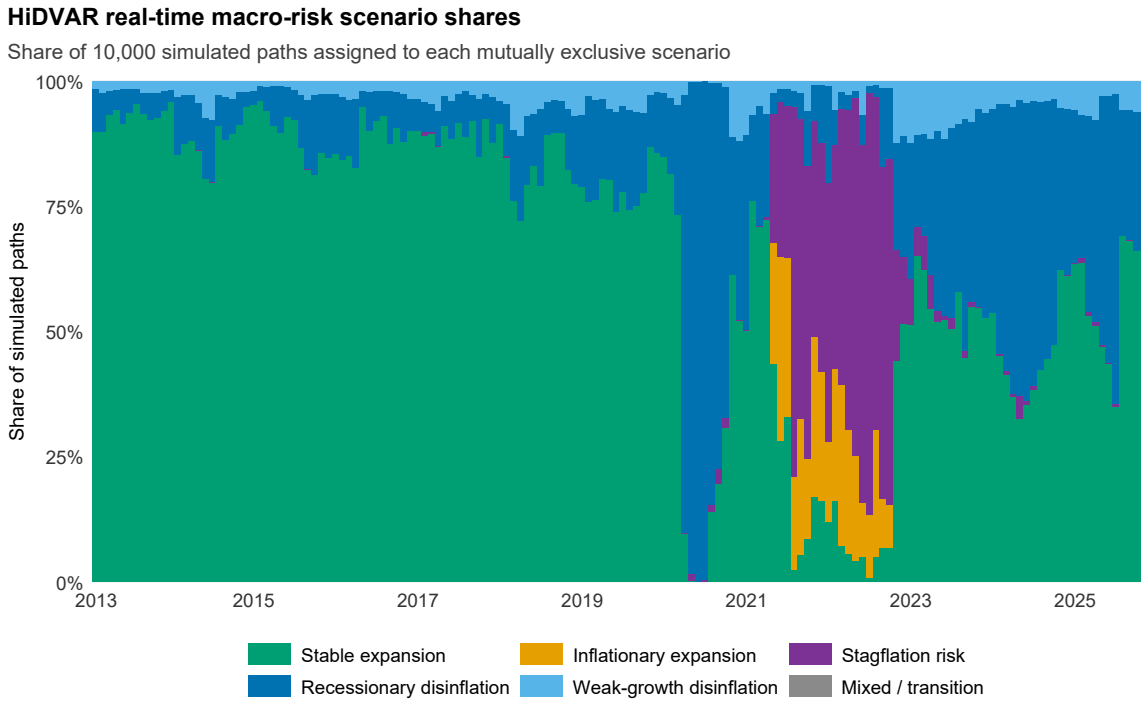


Figure 14: HiDVAR real-time macro-risk scenario shares.

Figure 14 shows that the model assigns most probability mass to stable expansion during the pre-pandemic period. From 2013 to 2019, the green area dominates the figure, while recessionary and weak-growth disinflation risks remain comparatively small. This is reassuring, since this period was characterized by moderate inflation and a relatively stable real economy.

The picture changes sharply around the time of the pandemic. The model assigns a large share of simulated paths to recessionary disinflation, reflecting the combination of severe real-side downside risk and initially limited persistent inflation pressure. This is exactly the type of episode in which a scenario classification is more informative than a single recession probability: the dominant risk was not merely weak growth, but weak growth in a disinflationary environment.

During 2021 and 2022, the model shifts from benign expansion toward inflationary and stagflation-risk scenarios. The orange area captures paths in which inflation pressure is persistent but real activity remains sufficiently robust. The purple area captures the more problematic policy configuration: persistent inflation pressure combined with weak growth, recession risk, or labor-market stress. The large purple share in this period indicates that the model did not interpret the inflation surge as a purely demand-driven boom. Instead, it placed substantial probability on a policy dilemma in which inflation remains elevated while real-side conditions deteriorate.

From 2023 onward, stagflation risk recedes markedly. The dominant risks become stable expansion and recessionary disinflation. This is consistent with a soft-landing interpretation, but not an unconditional one: the model still assigns non-negligible mass to hard-landing or weak-growth outcomes. In other words, the post-2022 forecast distribution is less about inflation remaining uncontrolled and more about the real costs of disinflation.

Overall, the scenario shares provide a compact summary of the model's distributional information. They show that HiDVAR distinguishes between very different macroeconomic risk configurations: benign expansions, inflationary booms, stagflationary policy dilemmas, and disinflationary downturns. This is more informative than reporting recession and stagflation probabilities separately, because it preserves the joint nature of the forecast distribution and makes clear whether inflation risk and real-side risk occur jointly or in isolation. Of course, this scenario analysis is possible with all of the models in use here, but it is substantially more informative with a single integrated, multivariate framework like HiDVAR or a BVAR-SV than with a combination of univariate approaches.

6.1.2 Inflation Tail-Risk Signals

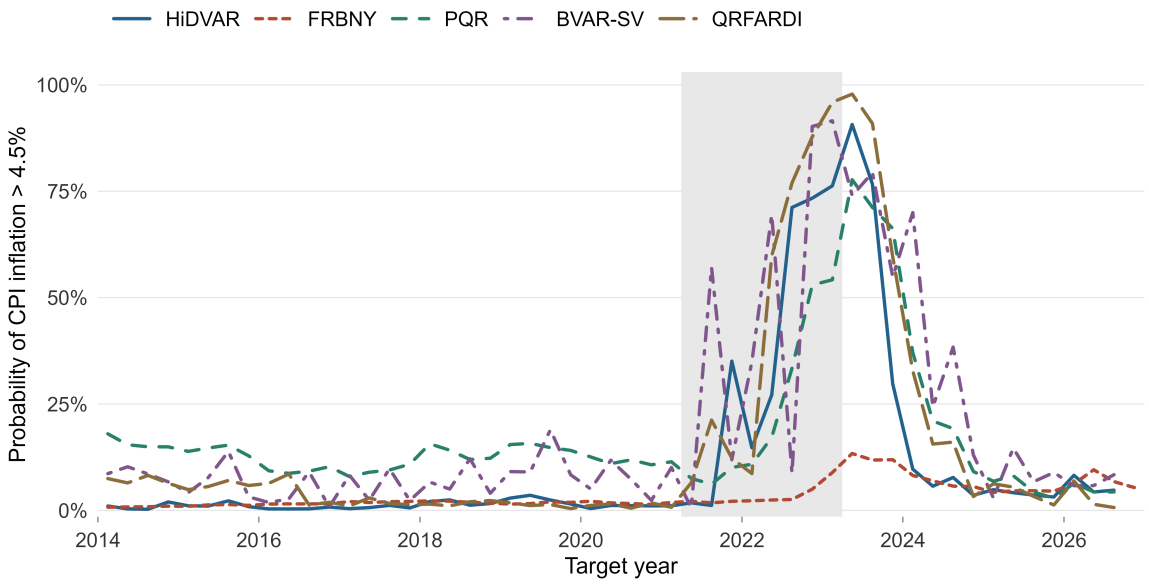


Figure 15: Predicted Probabilities of Inflation surpassing 4.5%

Note: Shaded region depicts period where actual inflation exceeded 4.5%.

Again, more event-oriented, I also document the probabilities of CPI inflation surpassing 4.5% for all models, expressed by the ratio of draws predicting high inflation compared to all draws in Figure 15. CPI inflation above 4.5% corresponds approximately to the upper tail of

the post-1990 US inflation distribution, with the 90th percentile of year-on-year CPI inflation equal to about 4.53%. It becomes even clearer that the FRBNY model and the professional forecasters did not see a proper inflation surge coming, rarely predicting high inflation for the future, even during 2022, peaking at only 16%. The HiDVAR, on the other hand, began putting a larger mass on a surge in prices already in early 2021, far earlier than professional forecasters or the FRBNY model. Also, it does so with much greater certainty. This can be shown with the Brier Score (Brier, 1950), which is a measure of event-probability accuracy, for the implied high-inflation probabilities by each model. Here, HiDVAR performs best,

Table 9: Brier Score for CPI Inflation >4.5%, N=132

	FRBNY	HiDVAR	PQR	BVAR-SV	QRFARDI
Brier Score	0.128	**0.078**	0.123	0.104	0.097

reducing the score by 39% over the FRBNY, or 20% over QRFARDI.

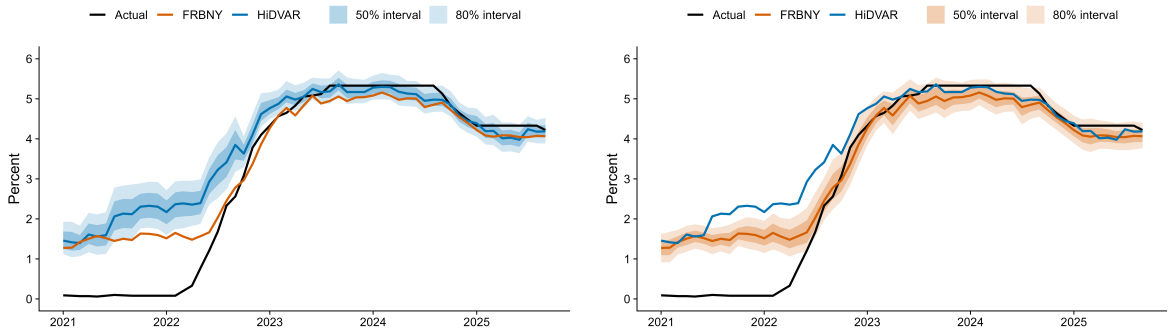
6.2 Policy-Stance Implications

In addition to the outcome variables themselves, one can also construct a policy Taylor-type rule using the forecasts to get a forward-looking policy rate distribution. Due to the Federal Reserve's focus on unemployment rather than output (Knotek, Verbrugge, Garciga, Treanor, & Zaman, 2016), I choose a Balanced Approach-type rule with forecasts from the models in place of the standard contemporary observations. Specifically, this forecast-based rule, parametrized analogously to Taylor (1993) can be expressed as:

$$i_t = \rho i_{t-1} + (1-\rho)(r_t^* + \pi^j_{t+h|t} + \phi_\pi(\pi^j_{t+h|t} - \pi^*) - \phi_u(u^j_{t+h|t} - u^*)) \quad j \in \{\text{HiDVAR, FRBNY}\} \tag{6}$$

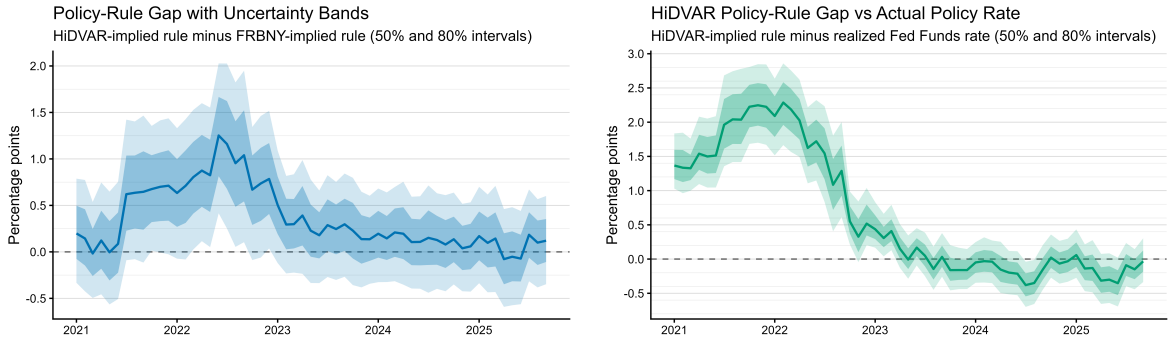
where I choose $\rho = 0.85$, r_t^* is equal to the Holston, Laubach, and Williams (2023) estimates in real-time, $\phi_\pi = 0.5$, $\phi_u = 1$, $\pi^* = 2\%$, u^* is equal to the noncyclical rate of unemployment , and the inflation and unemployment forecasts can be taken from the HiDVAR and FRBNY models.

Looking at the resulting rules post-2020 in Figures 16 and 17, one can see that both models prescribe higher interest rates from the start, something that is well-documented by Tatar and Wieland (2024), here mostly due to a positive r^*. During 2021, the HiDVAR began to recommend higher policy rates due to incoming inflation, whereas the FRBNY stayed lower. A positive surprise is the good fit that both rules provide from 2023 onwards, hinting that the OaR forecasts may have been policy-relevant, thus strengthening the use case for the HiDVAR. Moreover, this is also a strong application as the distributions of the federal funds rate are now also non-normal, with more upside risk in 2021 depicted by the HiDVAR.



(a) **(b)**

Figure 16: Forecast-Based policy rule with 50% and 80% density bands for HiDVAR (a) and FRBNY (b) against actual policy rate



(a) HiDVAR-implied against FRBNY-implied rule **(b)** HiDVAR implied Rule against actual Federal Funds

Figure 17: HiDVAR-implied Rule minus FRBNY and actual

7 Conclusion

This paper develops a HiDVAR, a real-time, data-rich framework for forecasting the joint predictive distribution of GDP growth, inflation, and unemployment without imposing parametric density restrictions. The key practical contribution is an operational density engine that allows the shape of forecast uncertainty, like skewness, tail thickness, and modality, to adjust endogenously to macro conditions, while maintaining multivariate and intertemporal coherence through a triangular factorization and copula-based path simulation.

Empirically, HiDVAR delivers its largest gains where policy risk management is most sensitive: inflation tail risk. In real-time out-of-sample evaluation from 2013 to 2025, HiDVAR materially improves probabilistic accuracy relative to the FRBNY Outlook-at-Risk densities and standard benchmarks for CPI inflation, with smaller and more comparable differences for GDP growth and unemployment. In real time, HiDVAR reduces CPI CRPS by 27.7% and the joint energy score by 21.8% relative to FRBNY OaR; the ES gain is weakly significant in HAC-adjusted DM tests (p=0.075). Importantly, during the post-pandemic inflation surge, the FRBNY densities assign implausibly low probability mass to persistent high inflation outcomes, whereas HiDVAR shifts substantial probability to the upper tail, yielding better-calibrated tail probabilities and more informative risk signals. The September 2021 case illustrates the mechanism behind the 'behind the curve' diagnosis: the FRBNY densities placed current inflation at an extreme percentile and assigned implausibly little mass to persistent high inflation, whereas HiDVAR shifted substantial probability into the upper tail and would have flagged elevated inflation risk materially earlier. HiDVAR also outperforms QRFARDI significantly for inflation, and weakly significant for the energy score.

The distributional forecasts also enable policy-relevant 'dashboard' objects that are not available from point forecasts alone. First, HiDVAR supports macro-risk scenario probabilities computed from coherent simulated paths. Second, translating the predictive distributions into rule-implied policy-rate prescriptions illustrates how density-based risk assessment can be mapped into interpretable measures of policy stance. These applications are intended as monitoring tools rather than estimates of the policy reaction function; nevertheless, they highlight that density restrictions and reliance on forecast anchors can systematically understate one-sided inflation risks in turbulent periods.

The results should be interpreted in light of several design choices. First, the framework is computationally intensive because it combines threshold-by-threshold distributional estimation with hyperparameter tuning and multi-horizon simulation, which limits the breadth of robustness checks that can be performed in routine real-time operation. Second, the strongest gains are concentrated in inflation and in the post-pandemic inflation episode; improvements for GDP growth and unemployment are smaller and, in some metrics, worse than benchmarks. Third, while the copula-based dependence layer ensures tractable path-consistent simulation, the Gaussian copula and structured dependence assumptions may miss richer forms of tail dependence or time variation in dependence. Fourth, calibration remains imperfect at extreme quantiles and differs across variables, indicating that tail-focused calibration diagnostics should remain a central part of evaluation.

These limitations do not overturn the paper's main contribution, but they define its scope: HiDVAR is best viewed as a risk-monitoring and density-forecasting framework rather than a structural model of policy transmission or a universally superior forecasting engine across all variables and episodes. Future work should therefore prioritize broader dependence-robustness exercises (alternative copula families and stress tests), applications to longer real-time out-of-sample time series, and extensions that improve computational scalability without sacrificing distributional flexibility. In a nutshell, more flexible density construction can materially improve real-time inflation risk monitoring in episodes where policy mistakes are driven less by the conditional mean than by mismeasured tail risk.

Bibliography

Adams, P. A., Adrian, T., Boyarchenko, N., & Giannone, D. (2021). Forecasting macroeconomic risks. *International Journal of Forecasting*, *37*(3), 1173–1191. Retrieved from https://ideas.repec.org/a/eee/intfor/v37y2021i3p1173-1191.html doi: 10.1016/j.ijforecast.2021.01.003

Adrian, T., Boyarchenko, N., & Giannone, D. (2019, April). Vulnerable growth. *American Economic Review*, *109*(4), 1263–1289. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.20161923 doi: 10.1257/aer.20161923

Adrian, T., Boyarchenko, N., & Giannone, D. (2021). Multimodality in macrofinancial dynamics. *International Economic Review*, *62*(2), 861–886. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/iere.12501 doi: 10.1111/iere.12501

Arnold, S., Walz, E.-M., Ziegel, J., & Gneiting, T. (2024). Decompositions of the mean continuous ranked probability score. *Electronic Journal of Statistics*, *18*(2), 4992–5044. doi: 10.1214/24-EJS2316

Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, *70*(1), 191–221. Retrieved from https://doi.org/10.1111/1468-0262.00273 doi: 10.1111/1468-0262.00273

Bai, J., & Ng, S. (2021). Matrix completion, counterfactuals, and factor analysis of missing data. *Journal of the American Statistical Association*, *116*(536), 1746–1763. Retrieved from https://doi.org/10.1080/01621459.2021.1967163 doi: 10.1080/01621459.2021.1967163

Bańbura, M., Giannone, D., Modugno, M., & Reichlin, L. (2013). Now-casting and the real-time data flow. In G. Elliott & A. Timmermann (Eds.), *Handbook of economic forecasting* (Vol. 2, pp. 195–237). Elsevier. Retrieved from https://doi.org/10.1016/B978-0-444-53683-9.00004-9 doi: 10.1016/B978-0-444-53683-9.00004-9

Bańbura, M., Giannone, D., & Reichlin, L. (2010). Large bayesian vector autoregressions. *Journal of Applied Econometrics*, *25*(1), 71–92. doi: 10.1002/jae.1137

Bańbura, M., & Modugno, M. (2014). Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics*, *29*(1), 133–160. doi: 10.1002/jae.2306

Belloni, A., & Chernozhukov, V. (2011). ℓ_1-penalized quantile regression in high-dimensional sparse models. *The Annals of Statistics*, *39*(1), 82–130. doi: 10.1214/10-AOS827

Bernanke, B. S. (2024, April). *Forecasting for monetary policy making and communication at the Bank of England: A review* (Independent Evaluation Office Report). Bank of England. Retrieved from https://www.bankofengland.co.uk/-/media/boe/files/independent-evaluation-office/2024/forecasting-for-monetary-policy-making-and-communication-at-the-bank-of-england-a-review.pdf

Bernanke, B. S., Boivin, J., & Eliasz, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly Journal of Economics*, *120*(1), 387–422. doi: 10.1162/0033553053327452

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, *78*(1), 1–3. doi: 10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2

Britton, E., Fisher, P., & Whitley, J. (1998). The inflation report projections: understanding the fan chart. *Bank of England Quarterly Bulletin*, *38*(1), 30–37.

Cahan, E., Bai, J., & Ng, S. (2023). Factor-based imputation of missing values and covariances in panel data of large dimensions. *Journal of Econometrics*, *233*(1), 113–131. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304407622000215 doi: 10.1016/j.jeconom.2022.01.006

Chavleishvili, S., & Manganelli, S. (2024). Forecasting and stress testing with quantile vector autoregression. *Journal of Applied Econometrics*, *39*(1), 66–85. doi: 10.1002/jae.3009

Chernozhukov, V., Fernandez-Val, I., & Galichon, A. (2009). Improving point and interval estimators of monotone functions by rearrangement. *Biometrika*, *96*(3), 559–575. Retrieved 2025-08-15, from http://www.jstor.org/stable/27798848 doi: 10.1093/biomet/asp030

Chernozhukov, V., Fernández-Val, I., & Galichon, A. (2010). Quantile and probability curves without crossing. *Econometrica*, *78*(3), 1093–1125. doi: 10.3982/ECTA7880

Chernozhukov, V., Fernández-Val, I., & Galichon, A. (2010). Quantile and probability curves without crossing. *Econometrica*, *78*(3), 1093–1125. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA7880 doi: 10.3982/ECTA7880

Chi, T.-C., Fan, T.-H., Ghigliazza, R. M., Giannone, D., & Wang, Z. K. (2025). *Macroeconomic forecasting and machine learning* (CEPR Discussion Paper No. 20727). Paris & London: Centre for Economic Policy Research. Retrieved from https://cepr.org/publications/dp20727

Clark, T. E. (2011). Real-time density forecasts from bayesian vector autoregressions with stochastic volatility. *Journal of Business & Economic Statistics*, *29*(3), 327–341. doi: 10.1198/jbes.2010.09248

Clements, M. P. (2004). Evaluating the bank of england density forecasts of inflation. *The Economic Journal*, *114*(498), 844–866. doi: 10.1111/j.1468-0297.2004.00246.x

Corradi, V., & Llorens-Terrazas, J. (2026). Monitoring joint tail risks: An application to growth and inflation. *Journal of Econometrics*, *255*, 106237. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304407626000588 doi: 10.1016/j.jeconom.2026.106237

Croushore, D., & Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, *105*(1), 111–130. doi: 10.1016/S0304-4076(01)00072-0

Demarta, S., & McNeil, A. J. (2005). The *t* copula and related copulas. *International Statistical Review*, *73*(1), 111–129. doi: 10.1111/j.1751-5823.2005.tb00254.x

Diebold, F. X., Gunther, T. A., & Tay, A. S. (1998). Evaluating density forecasts with applications to financial risk management. *International Economic Review*, *39*(4), 863–883. doi: 10.2307/2527342

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*(3), 253–263. Retrieved 2026-02-20, from http://www.jstor.org/stable/1392185 doi: 10.1080/07350015.1995.10524599

Doz, C., Giannone, D., & Reichlin, L. (2011). A two-step estimator for large approximate dynamic factor models based on kalman filtering. *Journal of Econometrics*, *164*(1), 188–205. Retrieved from https://www.sciencedirect.com/science/article/pii/S030440761100039X (Annals Issue on Forecasting) doi: 10.1016/j.jeconom.2011.02.012

Doz, C., Giannone, D., & Reichlin, L. (2012). A quasi–maximum likelihood approach for large, approximate dynamic factor models. *Review of Economics and Statistics*, *94*(4), 1014–1024. doi: 10.1162/REST_a_00225

Federal Reserve Bank of New York. (2023, May). *Technical background: "outlook-at-risk"* (Tech. Rep.). Federal Reserve Bank of New York, Research: Macrofinance Department. Retrieved from https://www.newyorkfed.org/medialibrary/Research/Interactives/outlook-at-risk/technicalnote (Accessed: 2025-11-10)

Foresi, S., & Peracchi, F. (1995). The conditional distribution of excess returns: An empirical analysis. *Econometrica*, *63*(2), 451–466. doi: 10.2307/2171955

Galí, J. (2015). *Hysteresis and the european unemployment problem revisited* (Working Paper No. 21430). National Bureau of Economic Research. Retrieved from https://www.nber.org/papers/w21430 doi: 10.3386/w21430

Giannone, D., Lenza, M., & Primiceri, G. E. (2015, 05). Prior selection for vector autoregressions. *The Review of Economics and Statistics*, *97*(2), 436–451. Retrieved from https://doi.org/10.1162/REST_a_00483 doi: 10.1162/REST_a_00483

Giannone, D., Reichlin, L., & Small, D. H. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, *55*(4), 665–676. (Dynamic factor model for nowcasting macro data) doi: 10.1016/j.jmoneco.2008.06.004

Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, *102*(477), 359–378. doi: 10.1198/016214506000001437

Gneiting, T., & Ranjan, R. (2011, July). Comparing density forecasts using threshold- and quantile-weighted scoring rules. *Journal of Business & Economic Statistics*, *29*(3), 411–422. Retrieved from https://ideas.repec.org/a/taf/jnlbes/v29y2011i3p411-422.html doi: 10.1198/jbes.2010.08110

Gneiting, T., & Resin, J. (2023). Regression diagnostics meets forecast evaluation: Conditional calibration, reliability diagrams, and coefficient of determination. *Electronic Journal of Statistics*, *17*(2), 3226–3286. doi: 10.1214/23-EJS2180

Goulet Coulombe, P., Leroux, M., Stevanovic, D., & Surprenant, S. (2022). How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*, *37*(5), 920–964. doi: 10.1002/jae.2910

Gourio, F., & Sarma, K. (2025, September). *Forecasting inflation during the pandemic: Who got it right?* (Chicago Fed Letter No. 513). Federal Reserve Bank of Chicago. Retrieved from https://www.chicagofed.org/publications/chicago-fed-letter/2025/513

Grishchenko, O. V., & Wilcox, L. (2024, May). *Tale about inflation tails* (Finance and Economics Discussion Series No. 2024-028). Board of Governors of the Federal Reserve System. Retrieved from https://www.federalreserve.gov/econres/feds/tale-about-inflation-tails.htm doi: 10.17016/FEDS.2024.028

Hannan, E. J., & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, *41*(2), 190–195. Retrieved from https://doi.org/10.1111/j.2517-6161.1979.tb01072.x doi: 10.1111/j.2517-6161.1979.tb01072.x

Henry, Ó. T., & Shields, K. (2004). Is there a unit root in inflation? *Journal of Macroeconomics*, *26*(3), 481–500. Retrieved from https://www.sciencedirect.com/science/article/pii/S016407040400031X doi: 10.1016/j.jmacro.2003.03.003

Holston, K., Laubach, T., & Williams, J. C. (2023, June). *Measuring the natural rate of interest after covid-19* (Staff Report No. 1063). Federal Reserve Bank of New York. Re-

trieved from https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr1063.pdf (Federal Reserve Bank of New York Staff Reports, No. 1063)

Jordà, Ò. (2005, March). Estimation and inference of impulse responses by local projections. *American Economic Review*, *95*(1), 161–182. Retrieved from https://www.aeaweb.org/articles?id=10.1257/0002828053828518 doi: 10.1257/0002828053828518

Kalman, R. E. (1960, 03). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, *82*(1), 35–45. Retrieved from https://doi.org/10.1115/1.3662552 doi: 10.1115/1.3662552

Kalman, R. E., & Bucy, R. S. (1961, 03). New results in linear filtering and prediction theory. *Journal of Basic Engineering*, *83*(1), 95–108. Retrieved from https://doi.org/10.1115/1.3658902 doi: 10.1115/1.3658902

Kilian, L., & Manganelli, S. (2008). The central banker as a risk manager: Estimating the federal reserve's preferences under greenspan. *Journal of Money, Credit and Banking*, *40*(6), 1103–1129. doi: 10.1111/j.1538-4616.2008.00150.x

Kim, H. H., & Swanson, N. R. (2014). Forecasting financial and macroeconomic variables using data reduction methods: New empirical evidence. *Journal of Econometrics*, *178*, 352–367. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304407613001978 (Recent Advances in Time Series Econometrics) doi: 10.1016/j.jeconom.2013.08.033

Knotek, E. S., II, Verbrugge, R. J., Garciga, C., Treanor, C., & Zaman, S. (2016, July). *Federal funds rates based on seven simple monetary policy rules* (Economic Commentary No. EC 2016-07). Federal Reserve Bank of Cleveland. Retrieved from https://www.clevelandfed.org/publications/economic-commentary/2016/ec-201607-federal-funds-rates-from-simple-policy-rules doi: 10.26509/frbc-ec-201607

Knüppel, M., Krüger, F., & Pohle, M.-O. (2022). *Score-Based Calibration Testing for Multivariate Forecast Distributions* (Discussion Paper No. 50/2022). Deutsche Bundesbank. Retrieved from https://www.bundesbank.de/en/publications/research/discussion-papers/score-based-calibration-testing-for-multivariate-forecast-distributions-900422

Koch, C., & Noureldin, D. (2024). How we missed the inflation surge: An anatomy of post-2020 inflation forecast errors. *Journal of Forecasting*, *43*(4), 852–870. doi: 10.1002/for.3088

Koenig, E. F., Dolmas, S., & Piger, J. (2003). The use and abuse of real-time data in economic forecasting. *The Review of Economics and Statistics*, *85*(3), 618–628. doi: 10.1162/003465303322369759

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica*, *46*(1), 33–50. Retrieved 2026-02-09, from http://www.jstor.org/stable/1913643 doi: 10.2307/1913643

Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, *4*, 83–91.

Korobilis, D. (2017). Quantile regression forecasts of inflation under model uncertainty. *International Journal of Forecasting*, *33*(1), 11–20. doi: 10.1016/j.ijforecast.2016.07.005

Korobilis, D., & Schröder, M. (2025). Monitoring multi-country macroeconomic risk: A quantile factor-augmented vector autoregressive (QFAVAR) approach. *Journal of*

Econometrics, *249*, 105730. doi: 10.1016/j.jeconom.2024.105730

Lenza, M., Moutachaker, I., & Paredes, J. (2025). Density forecasts of inflation: A quantile regression forest approach. *European Economic Review*, *178*, 105079. doi: 10.1016/j.euroecorev.2025.105079

López-Salido, D., & Loria, F. (2024). Inflation at risk. *Journal of Monetary Economics*, *145*, 103570. doi: 10.1016/j.jmoneco.2024.103570

Manzan, S. (2015). Forecasting the distribution of economic variables in a data-rich environment. *Journal of Business & Economic Statistics*, *33*(1), 144–164. doi: 10.1080/07350015.2014.937436

Manzan, S., & Zerom, D. (2013). Are macroeconomic variables useful for forecasting the distribution of U.S. inflation? *International Journal of Forecasting*, *29*(3), 469–478. doi: 10.1016/j.ijforecast.2013.01.005

McCracken, M. W., & Ng, S. (2016). FRED-MD: A Monthly Database for Macroeconomic Research. *Journal of Business & Economic Statistics*, *34*(4), 574–589. Retrieved from https://www.tandfonline.com/doi/full/10.1080/07350015.2015.1086655 doi: 10.1080/07350015.2015.1086655

McCracken, M. W., & Ng, S. (2021). FRED-QD: A Quarterly Database for Macroeconomic Research. *Federal Reserve Bank of St. Louis Review*, *103*(1), 1–44. Retrieved from https://www.stlouisfed.org/publications/review/2021/01/14/fred-qd-a-quarterly-database-for-macroeconomic-research doi: 10.20955/r.103.1-44

Meinshausen, N. (2006). Quantile regression forests. *Journal of Machine Learning Research*, *7*, 983–999. Retrieved from https://jmlr.org/papers/v7/meinshausen06a.html

Mitchell, J., Poon, A., & Zhu, D. (2024). Constructing density forecasts from quantile regressions: Multimodality in macrofinancial dynamics. *Journal of Applied Econometrics*, *39*(5), 790–812. doi: 10.1002/jae.3049

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, *47*(2), 527–556. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-2354.2006.00387.x doi: 10.1111/j.1468-2354.2006.00387.x

Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive–moving average models of unknown order. *Biometrika*, *71*(3), 599–607. doi: 10.1093/biomet/71.3.599

Scheuerer, M., & Hamill, T. M. (2015). Variogram-based proper scoring rules for probabilistic forecasts of multivariate quantities. *Monthly Weather Review*, *143*(4), 1321–1334. doi: 10.1175/mwr-d-14-00269.1

Schröder, M. (2026). Mixing it up: Inflation at risk. *Journal of Money, Credit and Banking*. (Advance online publication) doi: 10.1111/jmcb.70055

Sheather, S. J., & Jones, M. C. (1991, 07). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*, *53*(3), 683-690. Retrieved from https://doi.org/10.1111/j.2517-6161.1991.tb01857.x doi: 10.1111/j.2517-6161.1991.tb01857.x

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, *8*, 229–231.

Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, *19*(2), 279–281. doi: 10.1214/aoms/1177730256

Stock, J. H., & Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, *97*(460), 1167–1179. Retrieved 2025-12-08, from http://www.jstor.org/stable/3085839

doi: 10.1198/016214502388618960

Surprenant, S. (2025). *Quantile VARs and macroeconomic risk forecasting* (Staff Working Paper No. 2025-4). Bank of Canada. doi: 10.34989/swp-2025-4

Tatar, B., & Wieland, V. (2024). *Taylor rules and the inflation surge: The case of the fed* (CEPR Discussion Paper No. 18910). Retrieved from https://cepr.org/publications/dp18910

Taylor, J. B. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, *39*, 195–214. Retrieved from https://web.stanford.edu/~johntayl/Onlinepaperscombinedbyyear/1993/Discretion_versus_Policy_Rules_in_Practice.pdf doi: 10.1016/0167-2231(93)90009-L

Wallis, K. F. (1999). Asymmetric density forecasts of inflation and the bank of england's fan chart. *National Institute Economic Review*, *167*(1), 106–112. doi: 10.1177/002795019916700111

Wang, Y., Oka, T., & Zhu, D. (2026). Distributional vector autoregression: Eliciting macro and financial dependence. *Journal of Applied Econometrics*, *41*(4), 408–424. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.70049 doi: 10.1002/jae.70049

A Distributional Scoring Rules and Statistical Tests

Let F denote a predictive distribution on $\mathbb{R}^d$ and $y \in \mathbb{R}^d$ the realized outcome. A scoring rule $S(F, y)$ is *proper* if $\mathbb{E}_{Y\sim G}[S(G, Y)] \leq \mathbb{E}_{Y\sim G}[S(F, Y)]$ for all distributions F, G.

A.1 Univariate Scoring Rules

A.1.1 Continuous Ranked Probability Score (CRPS)

For univariate outcomes ($d = 1$), the CRPS (Gneiting & Raftery, 2007) is defined as

$$\mathrm{CRPS}(F, y) = \int_{-\infty}^{\infty} (F(z) - \mathbb{1}\{z \geq y\})^2 \quad , dz.$$

Equivalently,

$$\mathrm{CRPS}(F, y) = \mathbb{E}_F|X - y| - \tfrac{1}{2}\mathbb{E}_F|X - X'|,$$

where $X, X' \sim F$ are independent. The CRPS is a strictly proper scoring rule and generalizes the absolute error to probabilistic forecasts.

A.1.2 Logit Tail-Weighted CRPS

To emphasize forecast performance in the tails, I consider a weighted variant of the continuous ranked probability score. Let F denote a univariate predictive distribution and $y \in \mathbb{R}$ the realized outcome. For a nonnegative weight function $w(\cdot)$, the weighted CRPS is defined as

$$w\mathrm{CRPS}(F, y) = \int_{-\infty}^{\infty} w(z) \left(F(z) - \mathbf{1}\{z \geq y\}\right)^2 dz. \tag{7}$$

I use a smooth (logit) tail weight centered at a policy-relevant threshold u,

$$w(z) = \Lambda\Big(\frac{z-u}{s}\Big) = \frac{1}{1 + \exp\big(-\frac{z-u}{s}\big)}, \qquad s > 0, \tag{8}$$

which gradually increases the weight placed on the tails as z exceeds u. Given Monte Carlo draws $\{X^{(i)}\}_{i=1}^n$ from F, I evaluate wCRPS numerically on a grid $\{z_m\}_{m=1}^M$ using the empirical CDF $\widehat{F}(z) = n^{-1}\sum_{i=1}^n \mathbf{1}\{X^{(i)} \leq z\}$:

$$\widehat{w\mathrm{CRPS}}(F, y) = \sum_{m=1}^{M} w(z_m) \left(\widehat{F}(z_m) - \mathbf{1}\{z_m \geq y\}\right)^2 \Delta z. \tag{9}$$

As with the standard CRPS, lower values indicate better probabilistic accuracy, but with an explicit emphasis on tail performance.

A.1.3 Brier Score

The Brier score evaluates the accuracy of predictive probabilities for a specified event. For a univariate predictive distribution F and a threshold u, define the exceedance probability as

$$p_u = \Pr_F(Y > u) = 1 - F(u).$$

The corresponding Brier score is

$$\mathrm{BS}_u(F, y) = (p_u - \mathbb{1}\{y > u\})^2 . \tag{10}$$

It measures the squared difference between the predicted probability and the realized event indicator. Scores lie between zero and one, with lower values indicating greater predictive accuracy.

The Brier score is strictly proper for the probability of the specified binary event: its expected value is uniquely minimized by reporting the true event probability. At a fixed threshold, however, it does not identify the entire predictive distribution, since different distributions can imply the same exceedance probability.

For predictive draws $\{X^{(i)}\}_{i=1}^n$, the exceedance probability is estimated by

$$\widehat{p}_u = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{X^{(i)} > u\}. \tag{11}$$

For an evaluation sample of T forecasts and realizations, I report the average score,

$$\overline{\mathrm{BS}}_u = \frac{1}{T} \sum_{t=1}^{T} \left(\widehat{p}_{u,t} - \mathbb{1}\{y_t > u\}\right)^2 . \tag{12}$$

The Brier score provides a threshold-specific complement to the CRPS and its tail-weighted variant. In particular, the CRPS can be expressed as the integral of Brier scores over all thresholds,

$$\mathrm{CRPS}(F, y) = \int_{-\infty}^{\infty} \mathrm{BS}_u(F, y)\, du.$$

Whereas the CRPS assesses the full predictive distribution, the Brier score isolates forecast performance for a particular event, such as inflation exceeding 4.5%.

A.1.4 Interval Score

The interval score evaluates prediction intervals by combining their width with penalties for observations falling outside them. Let F denote a univariate predictive distribution and y the realized outcome. For a nominal coverage level $1 - \alpha$, define the central prediction interval by its lower and upper endpoints,

$$\ell_\alpha = F^{-1}(\alpha/2), \qquad u_\alpha = F^{-1}(1 - \alpha/2).$$

Following Gneiting and Raftery (2007), the interval score is

$$\mathrm{IS}_\alpha(F, y) =$$

$(\mathrm{u}_\alpha - \ell_\alpha) + \frac{2}{\alpha}(\ell_\alpha - y)\mathbf{1}y < \ell_\alpha + \frac{2}{\alpha}(y - u_\alpha)\mathbf{1}y > u_\alpha$.(13) The first term rewards narrow intervals, while the remaining terms penalize misses in proportion to their distance from the respective endpoint. If the realization lies inside the interval, the score equals its width. Lower scores therefore indicate a better balance between sharpness and coverage. The factor $2/\alpha$ increases the penalty for misses at higher nominal coverage levels.

The interval score is consistent for the pair of quantiles defining the central prediction interval: its expected value is minimized by the corresponding quantiles of the true distribution, uniquely when these quantiles are unique. At a fixed coverage level, however, it does not identify the entire predictive distribution, since distributions with identical interval endpoints receive identical scores.

In the empirical evaluation, I consider central 50% and 90% prediction intervals, corresponding to $\alpha = 0.50$ and $\alpha = 0.10$, respectively. Their endpoints are estimated from the predictive draws. For an evaluation sample of T forecasts and realizations, I report

$$\overline{\mathrm{IS}}_\alpha =$$

$1\frac{}{T\sum_{t=1}^{T}\mathrm{IS}_\alpha(F_t,y_t)}.(14)$ The table labels IS_{50} and IS_{90} refer to nominal coverage in percent. Comparisons across models are made separately at each coverage level. These scores complement the CRPS by assessing selected prediction intervals, although a lower interval score alone does not establish correct coverage or superior tail-event probability forecasts.

A.2 Multivariate Scoring Rules

A.2.1 Energy Score

For multivariate outcomes ($d \geq 1$), the energy score, introduced by Gneiting and Raftery (2007) is given by

$$\mathrm{ES}(F, y) = \mathbb{E}_F\|X - y\| - \tfrac{1}{2}\mathbb{E}_F\|X - X'\|,$$

where $X, X' \sim F$ are independent and $\|\cdot\|$ denotes the Euclidean norm. The energy score is a proper scoring rule for multivariate predictive distributions and reduces to the CRPS in the univariate case.

A.2.2 Variogram Score

To complement the energy score, I use the variogram score, which is more sensitive to misspecification of the dependence structure (Scheuerer & Hamill, 2015). For multivariate outcomes ($d \geq 2$), the variogram score of order $p > 0$ is defined as

$$\mathrm{VS}_p(F, y) = \sum_{i=1}^{d}\sum_{j=1}^{d} w_{ij}\left(|y_i - y_j|^p - \mathbb{E}_F|X_i - X_j|^p\right)^2,$$

where $X \sim F$ and $w_{ij} \geq 0$ are fixed weights. It compares observed pairwise differences with their predictive expectations, with lower scores indicating better predictive performance.

The variogram score is proper when the relevant moments exist, but is not strictly proper: different predictive distributions can imply the same expected pairwise differences. It therefore provides a complementary assessment of multivariate forecast quality rather than a complete evaluation of the joint distribution.

In my evaluation, I set $p = 0.5$ and use equal weights, $w_{ij} = 1$ for $i \neq j$. Observations and predictive draws are divided by the same fixed historical standard deviations used for the energy score. Predictive expectations are approximated by averages over the joint forecast draws.

A.3 Diebold–Mariano tests with CRPS and Energy Score

Predictive accuracy is compared using Diebold–Mariano (DM) tests (Diebold & Mariano, 1995) applied to losses obtained from proper scoring rules. For two competing density forecasts, denoted by models A and B, let

$$S\left(\widehat{F}_{t,h}^{(A)}, y_{t,h}\right) \quad \text{and} \quad S\left(\widehat{F}_{t,h}^{(B)}, y_{t,h}\right)$$

denote either the continuous ranked probability score (CRPS), the multivariate energy score (ES), or the variogram score (VS) at forecast horizon h. Both scoring rules are negatively oriented, such that lower values indicate greater predictive accuracy. The loss differential is defined as

$$d_t = S\left(\widehat{F}_{t,h}^{(A)}, y_{t,h}\right) - S\left(\widehat{F}_{t,h}^{(B)}, y_{t,h}\right), \qquad t = 1, \ldots, T. \tag{15}$$

Consequently, $d_t < 0$ indicates that model A produces a more accurate density forecast than model B.

The null hypothesis of equal predictive accuracy is

$$H_0 : \mathbb{E}[d_t] = 0. \tag{16}$$

Because the empirical analysis is concerned with whether model A outperforms the respective benchmark, inference is based on the directional alternative

$$H_1 : \mathbb{E}[d_t] < 0. \tag{17}$$

The direction of the alternative is therefore determined by the ordering of the models in the loss differential and is fixed prior to examining the test results.

The average loss differential is estimated using the constant-only regression

$$d_t = \mu + u_t. \tag{18}$$

The DM statistic is given by

$$\mathrm{DM} = \frac{\widehat{\mu}}{\widehat{\sigma}_{\mathrm{HAC}}(\widehat{\mu})}, \tag{19}$$

where $\widehat{\sigma}_{\mathrm{HAC}}(\widehat{\mu})$ is based on a heteroskedasticity- and autocorrelation-consistent Newey–West estimator.

Although the forecast horizon is four quarters, forecasts are produced and evaluated at a monthly frequency. Successive monthly vintages may therefore refer to the same quarterly target and incorporate overlapping information sets, data releases, and revisions. Serial dependence in the monthly loss differential can consequently extend beyond the conventional quarterly choice $h - 1 = 3$. To account for the twelve-month span associated with a four-quarter-ahead forecast when the evaluation sequence is indexed by monthly vintages, the HAC estimator is calculated using truncation lag

$$L = 11. \tag{20}$$

This allows the long-run variance estimator to incorporate autocovariances between loss differentials separated by up to eleven monthly forecast vintages.

The statistic is evaluated using the lower tail of a t distribution with $T - 1$ degrees of freedom. Accordingly, the reported one-sided p-value is

$$p_{\text{one-sided}} = \Pr(t_{T-1} \leq \text{DM}_{\text{obs}}) \,. \tag{21}$$

A sufficiently negative DM statistic therefore provides evidence that model A has a lower expected CRPS or energy score, and hence greater predictive accuracy, than model B.

A.4 Kolmogorov–Smirnov Test for PIT Uniformity

In the test by Kolmogorov (1933); Smirnov (1948), calibration is assessed via probability integral transforms (PITs). For a sequence of forecasts $\{F_t\}_{t=1}^T$ and realizations $\{y_t\}_{t=1}^T$, define

$$u_t = F_t(y_t), \qquad t = 1, \ldots, T. \tag{22}$$

Under correct calibration, $\{u_t\}$ should be approximately $\mathcal{U}(0, 1)$. Let $\widehat{H}_T(u) = T^{-1} \sum_{t=1}^T \mathbf{1}\{u_t \leq u\}$ denote the empirical CDF of the PITs. The Kolmogorov–Smirnov statistic for testing uniformity is

$$D_T = \sup_{u \in [0,1]} \left| \widehat{H}_T(u) - u \right| . \tag{23}$$

Large values of D_T (small p-values) indicate departures from uniformity and hence miscalibration of the predictive distribution.

A.5 Multivariate Score-Based Calibration Tests

To assess whether the joint predictive distributions are statistically consistent with the subsequent realizations, I employ the score-based multivariate calibration tests proposed by Knüppel et al. (2022). Let $\widehat{F}_t$ denote the joint predictive distribution of

$$\mathbf{Y}_t = \begin{pmatrix} \text{GDP growth}_t & \text{CPI inflation}_t & \text{unemployment}_t \end{pmatrix}' ,$$

and let $\mathbf{X}_t, \mathbf{X}_t^* \sim \widehat{F}_t$ denote independent draws from this distribution. The null hypothesis is multivariate auto-calibration,

$$H_0 : \mathcal{L}\left(\mathbf{Y}_t \mid \widehat{F}_t \right) = \widehat{F}_t, \tag{24}$$

meaning that the predictive distribution is correctly specified conditional on the information contained in the forecast itself.

The analysis uses the energy score,

$$\mathrm{ES}\left(\widehat{F}_t, \mathbf{y}_t\right) = \mathbb{E}_{\widehat{F}_t} \left\|\mathbf{X}_t - \mathbf{y}_t\right\| - \frac{1}{2}\mathbb{E}_{\widehat{F}_t} \left\|\mathbf{X}_t - \mathbf{X}_t^*\right\|, \tag{25}$$

where $\|\cdot\|$ denotes the Euclidean norm and lower values indicate better forecast performance.

The principal calibration test is the energy-score generalized Box transform (ES-GBT) test. For each forecast instance, the score-based probability integral transform is

$$U_{\mathrm{ES},t} = \Pr_{\widehat{F}_t}\left[\mathbb{E}_{\widehat{F}_t} \left\|\mathbf{X}_t - \mathbf{X}_t^*\right\| \leq \mathbb{E}_{\widehat{F}_t} \left\|\mathbf{X}_t - \mathbf{Y}_t\right\|\right]. \tag{26}$$

Given two separate forecast samples $\{\mathbf{X}_{t,i}\}_{i=1}^{J_0}$ and $\{\mathbf{X}_{t,j}^*\}_{j=1}^{J_1}$, this quantity is estimated by

$$\widehat{U}_{\mathrm{ES},t} = \frac{1}{J_1}\sum_{j=1}^{J_1} \mathbb{1}\left\{\frac{1}{J_0}\sum_{i=1}^{J_0}\left\|\mathbf{X}_{t,i} - \mathbf{X}_{t,j}^*\right\| \leq \frac{1}{J_0}\sum_{i=1}^{J_0}\left\|\mathbf{X}_{t,i} - \mathbf{Y}_t\right\|\right\}. \tag{27}$$

In the empirical implementation, a sample of up to $J = 5{,}000$ joint draws is divided into two non-overlapping subsamples. Under auto-calibration, $\widehat{U}_{\mathrm{ES},t}$ should be uniformly distributed on $[0, 1]$. I therefore test the first four uniform-moment restrictions,

$$\mathbb{E}\left[U_{\mathrm{ES},t}^k\right] = \frac{1}{k+1}, \qquad k = 1, \ldots, 4. \tag{28}$$

Defining

$$\mathbf{m}_t = \begin{pmatrix} U_{\mathrm{ES},t} - 1/2 \\ U_{\mathrm{ES},t}^2 - 1/3 \\ U_{\mathrm{ES},t}^3 - 1/4 \\ U_{\mathrm{ES},t}^4 - 1/5 \end{pmatrix}, \tag{29}$$

the corresponding Wald statistic is

$$\mathrm{GBT} = T\overline{\mathbf{m}}'\widehat{\mathbf{\Omega}}_{\mathbf{m}}^{-1}\overline{\mathbf{m}} \overset{a}{\sim} \chi_4^2, \tag{30}$$

where $\widehat{\mathbf{\Omega}}_{\mathbf{m}}$ is a HAC estimate of the long-run covariance matrix.

As a complementary diagnostic, I employ the energy score entropy test. It compares realized uncertainty with the uncertainty anticipated by the predictive distribution:

$$D_{\mathrm{ES},t} = \mathbb{E}_{\widehat{F}_t} \left\|\mathbf{X}_t - \mathbf{Y}_t\right\| - \mathbb{E}_{\widehat{F}_t} \left\|\mathbf{X}_t - \mathbf{X}_t^*\right\|. \tag{31}$$

Its split-sample estimator is

$$\widehat{D}_{\mathrm{ES},t} = \frac{1}{J_0}\sum_{i=1}^{J_0}\left\|\mathbf{X}_{t,i} - \mathbf{Y}_t\right\| - \frac{1}{J_0 J_1}\sum_{i=1}^{J_0}\sum_{j=1}^{J_1}\left\|\mathbf{X}_{t,i} - \mathbf{X}_{t,j}^*\right\|. \tag{32}$$

The entropy test evaluates

$$H_0 : \mathbb{E}\left[D_{\mathrm{ES},t}\right] = 0 \tag{33}$$

using

$$t_{\mathrm{ES}} = \frac{\overline{D}_{\mathrm{ES}}}{\sqrt{\widehat{\Omega}_D / T}}, \tag{34}$$

where $\widehat{\Omega}_D$ denotes the HAC long-run variance. A significantly positive value indicates overconfidence, since realized uncertainty exceeds the uncertainty anticipated by the model. A significantly negative value indicates underconfidence and therefore an excessively dispersed predictive distribution.

Because the forecasts are produced monthly for outcomes one year ahead, serial dependence induced by overlapping forecast windows is accounted for using Bartlett-kernel HAC covariance estimators with $L = 11$ lags.

B Parameters and restrictions used in the estimation

Throughout the estimation, several key parameters have to be estimated and calibrated. First, the empirical grid that decides the cutoff points of binary regressions is created by setting up an equidistant grid made up of the empirical percentiles of the dataset trimmed by 0.5% on both sides. To be certain to include all possible outcome values, the grid is augmented by a total of 5% of values that lie up to two standard deviations outside of the empirical grid. The grid is set to have tail-weights of 40%. The number of points in a grid per variable is set to 501. An example grid for QoQ values in May 2018 is provided below:

$$GDP_{Grid} = (-3.34, -3.17, -2.84, ..., 4.09, 4.61, 4.89) \tag{35}$$

$$CPI_{Grid} = (-1.48, -0.69, -0.67, ..., 3.22, 3.30, 3.59) \tag{36}$$

$$UR_{Grid} = (-1.12, -1.07, -0.99, ..., 1.96, 2.18, 2.29) \tag{37}$$

Also, for technical reasons, there is a maximum interval for all variables' outcomes. For GDP and inflation, these do not play a role, but for unemployment, there is a theoretical threshold of 2% applied, as lower values seem erroneously low given the historically observed rates.

For the model estimation, the tuned hyperparameters used in the elastic-net penalization are reported in Table 10; λ is reported separately per horizon. The possible grid includes values from 0 to 1 in 0.1 increments. As can be seen, the selected specifications generally imply relatively strong shrinkage, especially for GDP and unemployment, whereas it tends more towards ridge penalization for CPI inflation. The observed parameters are fairly standard in magnitude and do not show unusual behavior. It should be noted however, that for α, the training losses are quite similar across the grid such that values for α tend to change when using different tuning windows.

Table 10: Hyperparameter values for elastic-net penalization

Variable	α	$\lambda_{h=0}$	$\lambda_{h=1}$	$\lambda_{h=2}$	$\lambda_{h=3}$	$\lambda_{h=4}$
GDP	1.00	0.012	0.015	0.011	0.003	0.286
CPI	0.20	0.002	0.224	0.287	0.175	0.197
UR	0.90	0.003	0.011	0.018	0.011	0.028

C Copula Specification

C.1 Conditional Marginal Distributions

For each variable $j \in \{1, 2, 3\}$ and forecast horizon $h = 1, \ldots, H$, define the conditional cumulative distribution function (CDF)

$$F_{j,h}(y \mid X_t) = \Pr\left(Y^j_{t+h} \leq y \,\middle|\, X_t\right).$$

The CDF is approximated using a grid of thresholds $c_{j,h,1} < \cdots < c_{j,h,K}$. For each threshold, I estimate a penalized binary regression

$$\Pr\left(Y^j_{t+h} \leq c_{j,h,k} \mid X_t\right) = G\left(\beta^\top_{j,h,k} \tilde{X}_t\right),$$

where G is a logit link function and the parameters $\beta_{j,h,k}$ are obtained from an elastic-net penalized GLM with mixing parameter α_j and regularisation parameter $\lambda_{j,h}$, selected via external validation.

The resulting probability sequence is monotonized via rearrangement according to Chernozhukov, Fernandez-Val, and Galichon (2009). The corresponding probability integral transform (PIT) is

$$U^{j,h}_t = F_{j,h}\left(Y^j_{t+h} \mid X_t\right).$$

C.2 Gaussian Copula with Kronecker Correlation Structure

I chose this copula structure to yield tractability and parsimony given the relatively small sample size while ensuring positive definiteness and computational feasibility. Stack the PITs into the vector

$$U_t = \left(U^{1,1}_t, U^{2,1}_t, U^{3,1}_t, \ldots, U^{1,H}_t, U^{2,H}_t, U^{3,H}_t\right)^\top \in (0,1)^{3H}.$$

Let $\tilde{U}_t$ denote rank-based pseudo-observations. I estimate a Kronecker-structured correlation matrix

$$R = R_h \otimes R_v,$$

where:

- $R_v \in \mathbb{R}^{3\times 3}$ is the cross-sectional correlation matrix across variables. For variables j, k, Kendall's τ averaged over horizons is converted via

$$\rho_{jk} = \sin\left(\frac{\pi}{2}\,\bar{\tau}_{jk}\right).$$

- $R_h \in \mathbb{R}^{H\times H}$ models dependence across horizons using an AR(1) form,

$$(R_h)_{a,b} = \rho_h^{|a-b|}, \qquad \rho_h = \sin\left(\frac{\pi}{2}\bar{\tau}^{(1)}\right),$$

  where $\bar{\tau}^{(1)}$ is the average lag-1 Kendall τ across horizons and variables.

To ensure positive definiteness, I apply the shrinkage transformation

$$R \leftarrow (1-\eta)R + \eta I_{3H}, \qquad 0 < \eta \ll 1.$$

A Gaussian copula is then defined by

$$C_R(u_1, \ldots, u_{3H}) = \Phi_R\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_{3H})\right),$$

where Φ^{-1} is the univariate standard normal quantile and Φ_R is the CDF of a $3H$-dimensional normal distribution with correlation matrix R.

C.3 Simulation and Back–Transformation

To produce joint forecast paths, I draw

$$U^{(m)} \sim C_R, \qquad m = 1, \ldots, M,$$

and transform via the conditional marginal inverses:

$$Y_{t+h}^{j,(m)} = F_{j,h}^{-1}\left(U_{j,h}^{(m)} \mid X_t\right),$$

where the inverse CDF is computed numerically from the monotonized distribution estimate.

D Benchmarks

As benchmarks, the FRBNY OaR model, a Factor-Augmented BVAR with the baseline prior by Giannone et al. (2015) is included next to a more conventional Distributional Regression framework as well as a Quantile Random Forest with Autoregressive Diffusion Index (QRFARDI) with the same copula-structure to create multivariate output as the main model.

D.1 FRBNY *Outlook-at-Risk* benchmark (external density forecasts)

As a main benchmark, I use the *Outlook-at-Risk* (OaR) predictive distributions for real GDP growth, CPI inflation, and the unemployment rate described by Federal Reserve Bank of New York (2023) These forecasts are obtained directly from the FRBNY OaR release and are used as provided; I do not re-estimate the OaR model.

OaR constructs monthly conditional distributions using a two-step procedure based on quantile regressions and parametric smoothing. Let t index months and let $q(t)$ denote the corresponding calendar quarter. For horizon $h \in \{1, \dots, H\}$, define the target $y_{q(t)+h}$ as the h-quarter-ahead value (unemployment) or the annualized h-quarter average growth rate (GDP growth and CPI inflation). Let $\hat{y}^{BC}_{q(t)+h|t}$ be the consensus Blue Chip Economic Indicators (BCEI) forecast and $e^{BC}_{q(t)+h|t} \equiv y_{q(t)+h} - \hat{y}^{BC}_{q(t)+h|t}$ the associated forecast error.

For each horizon h and quantile level $\tau \in (0, 1)$, OaR estimates linear quantile regressions of forecast errors on conditioning information x_t observed in month t (including a constant and CISS as a financial-conditions indicator),

$$Q_{e,h}(\tau \mid x_t) = x_t^\top \beta_h(\tau).$$

Under the maintained assumption that the BCEI forecast and its forecast error are independent conditional on x_t, the implied conditional quantiles of the target are obtained by shifting the estimated error quantiles by the BCEI point forecast,

$$Q_{y,h}(\tau \mid x_t) = \hat{y}^{BC}_{q(t)+h|t} + Q_{e,h}(\tau \mid x_t).$$

Because the first step delivers quantile estimates on a finite grid of τ's, OaR fits a flexible four-parameter skew-t distribution (Azzalini–Capitanio) to these implied quantiles by minimizing the squared distance between the estimated quantiles and the model-implied quantiles. The fitted distribution then defines a full predictive CDF $\widehat{F}^{OaR}_{h,t}$, from which tail-risk measures (e.g. lower-tail quantiles) and predictive densities are computed.

The OaR distributions are updated each month using an expanding real-time information set, with periodic re-estimation on a fixed schedule to avoid look-ahead bias.

D.2 Penalized quantile-regression benchmark with Student-t copula

As an additional distributional benchmark, I estimate a factor-augmented penalized quantile-regression process (PQR) and combine the resulting conditional marginal distributions using a Student-t copula. The conditional quantile function is estimated directly on a grid of quantile levels and subsequently converted into a complete predictive distribution. This follows the nonparametric density-construction approach advocated by Mitchell et al. (2024), while the recursive ordering of the target variables is related to the quantile vector autoregression of Chavleishvili and Manganelli (2024).

Let

$$Y_t = \left(y_t^{(1)}, \dots, y_t^{(V)}\right)^\top$$

denote the vector of target variables and let f_t contain the factors constructed identically to the main specification. Defining $Z_t = (Y_t^\top, f_t^\top)^\top$, the common predictor vector available at forecast origin t is

$$x_t = \left(f_t^\top, Z_{t-1}^\top, \dots, Z_{t-L}^\top\right)^\top,$$

where L is the fixed lag order. Thus, the model uses the current factor estimates together with lagged target variables and lagged factors.

Separate direct forecasting equations are estimated for each horizon $h = 1, \dots, H$. To retain a multivariate structure, the target variables are ordered recursively. For variable j, define

$$d_{t,h}^{(j)} = \left(x_t^\top, y_{t+h}^{(1)}, \dots, y_{t+h}^{(j-1)}\right)^\top,$$

where the contemporaneous lower-ordered outcomes are omitted for $j = 1$. The conditional quantile function is specified as

$$Q_{y_{t+h}^{(j)}}\left(\tau \mid x_t, y_{t+h}^{(1)}, \dots, y_{t+h}^{(j-1)}\right) = \beta_{0,j,h}(\tau) + d_{t,h}^{(j)\top}\beta_{j,h}(\tau), \qquad \tau \in \mathcal{T}.$$

This recursive factorization allows the conditional relationship between the target variables to differ across the outcome distribution, as in the structural quantile framework of Chavleishvili and Manganelli (2024). At the same time, the use of direct equations permits the coefficients and selected predictors to vary across forecast horizons.

The quantile grid is

$$\mathcal{T} = \{0.025, 0.05, 0.075, 0.10, 0.15, \dots, 0.85, 0.90, 0.925, 0.95, 0.975\},$$

which places additional resolution in both tails. For each variable and horizon, the coefficient vectors are estimated by minimizing the quantile check loss subject to an elastic-net penalty:

$$\left\{\widehat{\beta}_{0,j,h}(\tau), \widehat{\beta}_{j,h}(\tau)\right\}_{\tau\in\mathcal{T}} = \arg\min_{\{\beta_0(\tau),\beta(\tau)\}} \frac{1}{|\mathcal{T}|N_{j,h}} \sum_{\tau\in\mathcal{T}} \sum_{t\in\mathcal{I}_{j,h}} \rho_\tau\left(y_{t+h}^{(j)} - \beta_0(\tau) - d_{t,h}^{(j)\top}\beta(\tau)\right) + \lambda_{j,h} \sum_{\tau\in\mathcal{T}} \left[\alpha_{j,h}\|\beta(\tau)\|_1 + \frac{1-\alpha_{j,h}}{2}\|\beta(\tau)\|_2^2\right],$$

where

$$\rho_\tau(e) = e\,(\tau - \mathbb{1}\{e < 0\})$$

is the quantile check loss. The elastic-net penalty permits both variable selection and shrinkage of correlated predictors and makes estimation feasible even when the predictor dimension is large relative to the available time-series sample. This use of regularization is closely related to the high-dimensional penalized quantile-regression framework of Belloni and Chernozhukov (2011).

The penalty parameters $\lambda_{j,h}$ and $\alpha_{j,h}$ are selected separately for every target variable and forecast horizon using a chronological validation block. The final portion of the estimation sample is held out, and each candidate specification is evaluated using a weighted average of the quantile losses,

$$\mathrm{QS}_{j,h}(\lambda,\alpha) = \sum_{k=1}^{|\mathcal{T}|} w_k \frac{1}{N_{\mathrm{val}}} \sum_{t\in\mathcal{I}_{\mathrm{val}}} \rho_{\tau_k}\left(y_{t+h}^{(j)} - \widehat{Q}_{j,h}(\tau_k \mid d_{t,h}^{(j)})\right),$$

where the weights w_k correspond to midpoint cells around the potentially unevenly spaced quantile levels. This criterion numerically approximates the integrated quantile score and is therefore proportional to the continuous ranked probability score. The selected hyperparameters satisfy

$$(\widehat{\lambda}_{j,h}, \widehat{\alpha}_{j,h}) = \arg\min_{\lambda,\alpha} \mathrm{QS}_{j,h}(\lambda,\alpha).$$

The resulting hyperparameters are reused within the forecast year and retuned annually, thereby limiting computational cost without using future observations.

Because quantile regressions are estimated separately across quantile levels, their raw predictions need not be monotone in τ. I therefore sort the predicted quantiles at each covariate value, implementing the monotone-rearrangement principle of Chernozhukov, Fernández-Val, and Galichon (2010). Let

$$\widehat{Q}_{j,h}^{\,r}(\tau \mid d)$$

denote the rearranged conditional quantile curve. Between adjacent fitted quantile levels, the curve is completed by linear interpolation. Outside the estimated range $[0.025, 0.975]$, each tail is extrapolated linearly in the logit of the quantile level:

$$\widehat{Q}^{r}_{j,h}(\tau \mid d) = a_{s,j,h}(d) + b_{s,j,h}(d) \log\left(\frac{\tau}{1-\tau}\right), \qquad s \in \{\text{lower}, \text{upper}\},$$

where the tail-specific coefficients are estimated from the four quantile points closest to the respective boundary. This yields a continuous quantile function over $(0, 1)$ without imposing symmetry, unimodality, or a particular parametric marginal distribution.

The marginal quantile processes are combined through a Student-t copula. Importantly, the dependence parameters are not estimated from in-sample fitted residuals. Instead, I construct expanding-window, blocked out-of-sample probability integral transforms,

$$u_{t,j,h} = \widehat{F}^{(-t)}_{j,h}\left(y^{(j)}_{t+h} \mid d^{(j)}_{t,h}\right),$$

where $\widehat{F}^{(-t)}_{j,h}$ is obtained by numerically inverting the rearranged quantile curve estimated using only observations preceding the corresponding evaluation block. The PITs are stacked across variables and horizons,

$$U_t = \left(u_{t,1,1}, \ldots, u_{t,V,1}, u_{t,1,2}, \ldots, u_{t,V,H}\right)^{\top},$$

and the most recent complete PIT vectors are used to estimate the copula.

To obtain a parsimonious dependence structure, the copula correlation matrix is specified as

$$R = R_H \otimes R_V,$$

where R_V captures dependence across target variables and R_H captures dependence across forecast horizons. Both components are estimated from Kendall rank correlations. For an elliptical copula, Kendall's τ is mapped into a linear correlation using

$$\rho = \sin\left(\frac{\pi}{2}\tau^{K}\right).$$

The resulting correlation matrices are mildly shrunk toward the identity matrix and projected onto the set of valid correlation matrices to ensure numerical stability. The degrees-of-freedom parameter ν is selected by maximizing the Student-t copula likelihood over a discrete candidate grid. Relative to a Gaussian copula, the Student-t copula permits symmetric tail dependence and therefore allows adverse or unusually favorable outcomes to occur jointly across variables and horizons (Demarta & McNeil, 2005).

For every forecast vintage, I simulate

$$U_n = \left(U_{n,1,1}, \ldots, U_{n,V,H}\right)^{\top} \sim C^{t}_{\widehat{\nu},\widehat{R}}, \qquad n = 1, \ldots, N,$$

from the fitted Student-t copula. Forecast draws are then generated recursively within each horizon:

$$\widetilde{y}^{(j)}_{n,h} = \widehat{Q}^{r}_{j,h}\left(U_{n,j,h} \mid x_T, \widetilde{y}^{(1)}_{n,h}, \ldots, \widetilde{y}^{(j-1)}_{n,h}\right).$$

Thus, the simulated lower-ordered target variables enter the conditional quantile functions of the higher-ordered variables. The copula additionally preserves residual dependence across variables and forecast horizons. I generate $N = 10{,}000$ joint forecast paths and apply the same back-transformation as in the main model to express GDP growth and inflation in year-on-year units and unemployment in levels.

Overall, the PQR-t benchmark combines four elements: high-dimensional elastic-net quantile regression, nonparametric reconstruction of the conditional marginals, recursive multivariate conditioning, and Student-t-copula dependence. It does not impose a skew-t or any other parametric family on the marginal predictive distributions. It can therefore accommodate asymmetric, heavy-tailed, and potentially multimodal forecast distributions while retaining coherent joint simulations across target variables and horizons.

D.3 Factor-augmented BVAR with stochastic volatility

As a benchmark, I estimate a factor-augmented Bayesian VAR with stochastic volatility (BVAR–SV) on the same stacked data vector as in the main model,

$$Z_t = (y_t^\top, f_t^\top)^\top,$$

where y_t contains the target variables and f_t denotes the factor block, constructed identically to the main specification. Following the two-step FAVAR approach of Bernanke, Boivin, and Eliasz (2005), the factors are first extracted from the large underlying information set, subsequently treated as observed, and included as endogenous variables alongside the target variables.

The conditional mean follows a VAR(p) with an intercept,

$$Z_t = c + \sum_{\ell=1}^{p} A_\ell Z_{t-\ell} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \Sigma_t),$$

where, in contrast to the homoskedastic BVAR, the innovation covariance matrix Σ_t varies over time. The use of stochastic volatility follows the BVAR density-forecasting literature, in which allowing innovation variances to evolve over time can substantially improve the calibration and accuracy of macroeconomic predictive distributions (Clark, 2011). The time-varying covariance matrix is decomposed as

$$\Sigma_t = B^{-1} H_t (B^{-1})^\top, \qquad H_t = \operatorname{diag}\left(e^{h_{1t}}, \ldots, e^{h_{nt}}\right),$$

where B is a lower-triangular contemporaneous-impact matrix and h_{it} is the log innovation variance associated with the i-th orthogonalized disturbance. Each log variance evolves according to a stationary autoregressive process,

$$h_{it} = \mu_i + \phi_i (h_{i,t-1} - \mu_i) + \sigma_i^h \eta_{it}, \qquad \eta_{it} \sim \mathcal{N}(0, 1),$$

with $|\phi_i| < 1$. This specification allows the magnitude and covariance structure of macroeconomic shocks to adjust over time, which is particularly relevant for a sample containing episodes of pronounced changes in macroeconomic volatility.

The prior for the VAR coefficients follows the hierarchical Minnesota-style shrinkage framework of Giannone et al. (2015). Conditional on the latent volatility paths and the remaining covariance parameters, the coefficient vector

$$\beta = \text{vec}(c, A_1, \ldots, A_p)$$

is assigned a Gaussian prior that shrinks the dynamics toward a parsimonious benchmark: the prior mean is centered on a random walk for the first own lag of persistent variables and on zero for other coefficients. Prior variances decrease with the lag order and depend on the relative scale of the variables. Their informativeness is governed by hyperparameters such as the overall tightness λ, lag decay α, and cross-variable scaling parameters ψ.

In the spirit of Giannone et al. (2015), these shrinkage hyperparameters are treated as unknown rather than fixed through ad hoc calibration. Because stochastic volatility breaks the Normal–inverse-Wishart conjugacy of the homoskedastic specification, the VAR coefficients, shrinkage hyperparameters, contemporaneous covariance parameters, and latent log-volatility processes are inferred jointly using MCMC. Conditional Gaussian updates are used where available, while non-conjugate parameter blocks are sampled using appropriate Metropolis–Hastings or state-space simulation steps. Priors on the stochastic-volatility parameters regularize the unconditional variance, persistence, and innovation variance of each log-volatility process.

Density forecasts are constructed from the posterior predictive distribution. For every retained posterior draw, future log-volatility states are first simulated from their transition equations. Conditional on these states, future VAR innovations are drawn from

$$\varepsilon_{T+h} \sim \mathcal{N}(0, \Sigma_{T+h}),$$

and the VAR is iterated forward to obtain a complete forecast path. The resulting draws therefore incorporate uncertainty about the VAR coefficients, the shrinkage and covariance parameters, future shocks, and the evolution of macroeconomic volatility. Finally, I retain the draws corresponding to the target variables and apply the same back-transformation as in the main model to report predictive distributions in year-on-year and level units.

D.4 Quantile random forest with autoregressive diffusion indices

As a nonlinear and data-rich machine-learning benchmark, I employ a Quantile Random Forest with Autoregressive Diffusion Indices (QRFARDI), following the forecasting framework popularized by Goulet Coulombe et al. (2022). The model combines the ability of random forests to capture nonlinearities and interactions with the dimension-reduction benefits of diffusion indices. For each target variable $v \in \{1, \ldots, V\}$ and forecast horizon $h \in \{1, \ldots, H\}$,

the predictor vector contains lags of the target variables together with the common factors extracted from the large macroeconomic information set:

$$\mathbf{z}_t = \left(\mathbf{y}_{t-1}^\top, \ldots, \mathbf{y}_{t-L}^\top, \mathbf{f}_t^\top, \ldots, \mathbf{f}_{t-L_f}^\top\right)^\top.$$

Separate direct models are estimated for each variable and horizon, thereby avoiding the accumulation of errors associated with iterated forecasts.

Quantile regression forests extend conventional random forests by retaining the conditional distribution of observations contained in the terminal nodes rather than only their conditional mean (Meinshausen, 2006). Let $w_{t,i}^{(v,h)}(\mathbf{z}_t)$ denote the forest weight assigned to historical observation i when forecasting variable v at horizon h. The estimated conditional cumulative distribution function is

$$\widehat{F}_{v,h}(y \mid \mathbf{z}_t) = \sum_{i=1}^{T-h} w_{t,i}^{(v,h)}(\mathbf{z}_t)\, \mathbb{1}\left\{y_{i+h}^{(v)} \leq y\right\},$$

and the corresponding conditional quantile forecast is

$$\widehat{Q}_{v,h}(\tau \mid \mathbf{z}_t) = \inf\left\{y : \widehat{F}_{v,h}(y \mid \mathbf{z}_t) \geq \tau\right\}, \qquad \tau \in \mathcal{T},$$

where $\mathcal{T}$ denotes a grid of quantile levels. Forest hyperparameters, including the number of candidate predictors considered at each split, the minimum terminal-node size, and the subsampling fraction, are selected using rolling time-series validation.

Because the estimated quantiles are available only on the finite grid $\mathcal{T}$, I convert them into a smooth predictive density by fitting a four-parameter skew-t distribution. For each variable, horizon, and forecast origin, its location, scale, skewness, and tail-thickness parameters are chosen to minimize the discrepancy between the forest quantiles and the corresponding skew-t quantiles:

$$\widehat{\boldsymbol{\theta}}_{v,h,t} = \arg\min_{\boldsymbol{\theta}} \sum_{\tau \in \mathcal{T}} \omega_\tau \left[Q_{\mathrm{ST}}(\tau; \boldsymbol{\theta}) - \widehat{Q}_{v,h}(\tau \mid \mathbf{z}_t)\right]^2,$$

where $\boldsymbol{\theta} = (\mu, \sigma, \alpha, \nu)^\top$ and ω_τ denotes the weight assigned to quantile level τ. The fitted distribution defines a complete predictive CDF from which forecast draws and density-based risk measures can be obtained.

Finally, the horizon- and variable-specific marginal distributions are connected using the same Gaussian copula and cross-horizon dependence structure as in the HiDVAR. Copula draws are transformed through the fitted skew-t inverse CDFs, producing multivariate and serially coherent forecast paths. QRFARDI therefore provides a demanding benchmark that is nonlinear, data-rich, and capable of representing asymmetric and heavy-tailed uncertainty. Unlike HiDVAR, however, its final predictive density remains restricted to the selected skew-t family and is consequently generally unable to preserve multimodality or other irregular features contained in the underlying forest quantiles.

E Stationarity of Variables

As my target is year-on-year growth, I face a dilemma about whether my variables are stationary, given this target. Henry and Shields (2004) already discussed that there may be unit roots in YoY inflation, or Galí (2015) for the level unemployment rate. For checking my US series, I use an Augmented Dickey-Fuller test (Said & Dickey, 1984) with k=4 to test the stationarity of YoY GDP, YoY CPI, and the level unemployment rate since 1960. While GDP turns out to be stationary already, the other two cannot be safely determined. To keep variables and targets aligned, however, I choose GDP and CPI targets in alignment. Thus, I transform GDP and CPI into QoQ percentage changes and unemployment into first differences. Now, all three variables are clearly stationary, aligned, and ready to work with. This means that for getting YoY forecasts, I simply add up 4 QoQ growth rates for GDP and CPI, and add the 4 changes in unemployment to the original level of unemployment.

Table 11: ADF unit root tests before and after transformation

	YoY series			QoQ series		
Variable	**ADF**	**p-val.**	**Tentative Interpretation**	**ADF**	**p-val.**	**Tentative Interpretation**
GDP growth	−3.85	0.020	Stationary (YoY)	−5.48	<0.01	Stationary (QoQ)
CPI inflation	−2.49	0.373	Non-stationary (YoY)	−4.02	0.011	Stationary (QoQ)
Unemployment rate	−2.60	0.327	Non-stationary (level)	−5.42	<0.01	Stationary (Δ)

Notes: The ADF test has the null hypothesis of a unit root (points to non-stationarity). Decisions use a 5% significance level. "Transformed" uses QoQ for GDP and CPI and first differences for unemployment.

F Robustness check with a different triangular ordering

As the triangular ordering is debatable, especially concerning the order of CPI inflation and unemployment, I present alternative orderings here. The baseline ordering is $GDP \rightarrow UR \rightarrow CPI$.

Table 12 shows that the causal ordering only has a small effect on the overall results, but does not inherently transform those.

	Univariate (CRPS)			**Multivariate**	
Ordering	GDP	CPI	UR	ES	VS
Baseline	0.670	0.729	0.315	0.813	0.972
GDP→CPI→UR	0.670	0.732	0.319	0.816	0.967
UR→GDP→CPI	0.639	0.733	0.363	0.822	1.003
CPI→UR→GDP	0.632	0.728	0.355	0.815	0.979

Table 12: Evaluation Scores under Alternative Variable Orderings

G Data

G.1 Introduction

For my forecasting projects, I wanted a quarterly, real-time machine learning ready dataset for use in macroeconomic projections. The natural choice here is the FRED-QD dataset by McCracken and Ng (2021). It comprises around 250 variables and contains quarterly data dating back to 1959. However, its real-time construction so far only dates back to May 2018, which may be too short for proper out-of-sample analysis.

For this, I first replicate the FRED-QD dataset as a whole. This is done from two sources, ALFRED and the monthly counterpart of FRED-QD, FRED-MD (McCracken & Ng, 2016). I try to follow the construction of the original paper (McCracken & Ng, 2021). I download all relevant time series with the 1st of the next month as a vintage date with the ALFRED API in R. That means that the vintage date for the June 2017 dataset is the 1st of July 2017 to align with the publication schedule of the original authors. To help with many of the spliced series and the similarity of the two datasets, I also import variables wherever possible from FRED-MD, which dates back to 1999 in real-time. With that, I can date back the dataset to January 2001. When comparing my results and the descriptions to the original and updated appendix, I noticed some inconsistencies in the latter tables highlighted in red.

Moreover, I thought that the dataset could even be enlarged with some common time series, like the overall working population, inflation expectations, or more granular survey responses from the SLOOS Bank Officer Survey. In total, 15 variables are added over the full sample; see tables 13 until 29 where new variables are highlighted in blue.

However, due to the unavailability of series at earlier vintages, the number of variables varies over time, ranging from 190 in January 2001 to 263 in May 2018, when the standard FRED-QD starts. Sometimes, it is possible to replace a variable or recalculate it manually; otherwise, the series will be dropped. For details, see tables 30 and 31.

Furthermore, due to the quarterly structure, the data has significant lags, even for variables only available in monthly frequencies. These current-quarter observations can be used to nowcast the dataset alongside improved missing value imputation, all done with mixed frequency dynamic factor models as shown below.

In Figure 18, I document the trajectory of the number of variables and time series observations. The biggest cut of variables occurred in 2013 with the Introduction of more granular NIPA data by the BEA. Overall, there are usually more variables than time-series observations, although the dataset remains approximately square over the full sample.

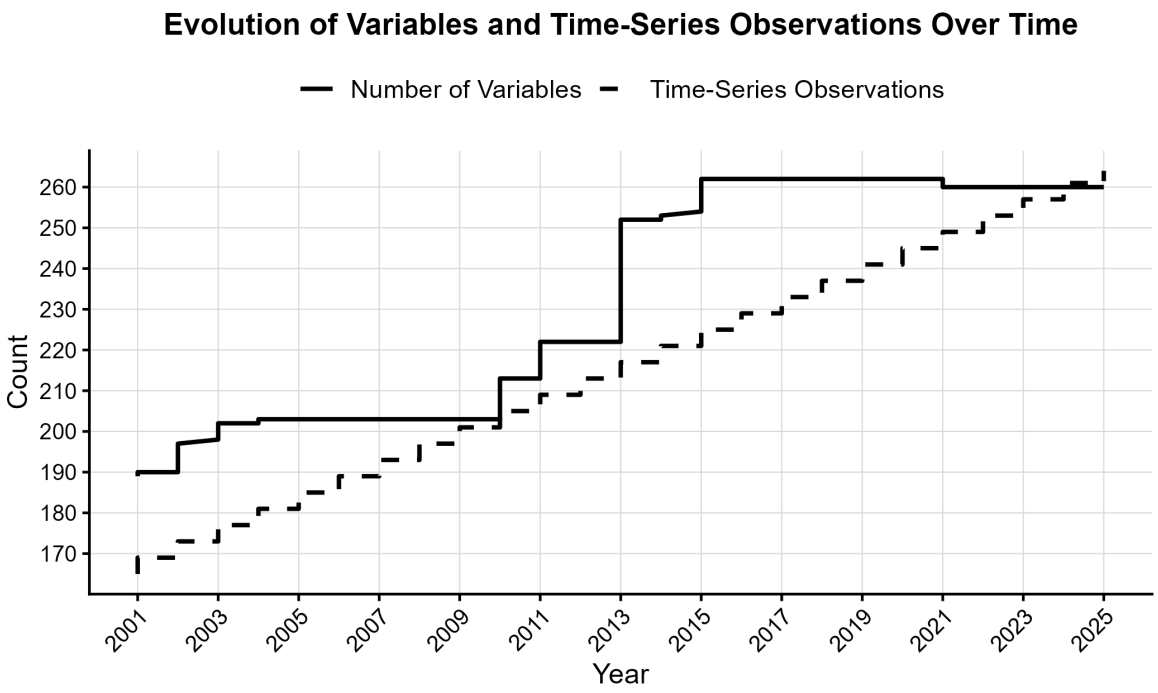


Figure 18: Evolution of Dataset Dimension

G.2 Nowcasting and Missing-Value Imputation

G.2.1 Motivation and real-time timing

The baseline quarterly panel exhibits a publication lag, while roughly two-thirds of the information set is observed at a monthly frequency. To reduce this lag in real time, I augment the quarterly dataset with nowcasts for the *second* and *third* month of each quarter, when sufficiently many current-quarter monthly releases are available. For the first month of a quarter (e.g., January, April, July, October), this is typically not feasible because only a little representative current-quarter information has been published.

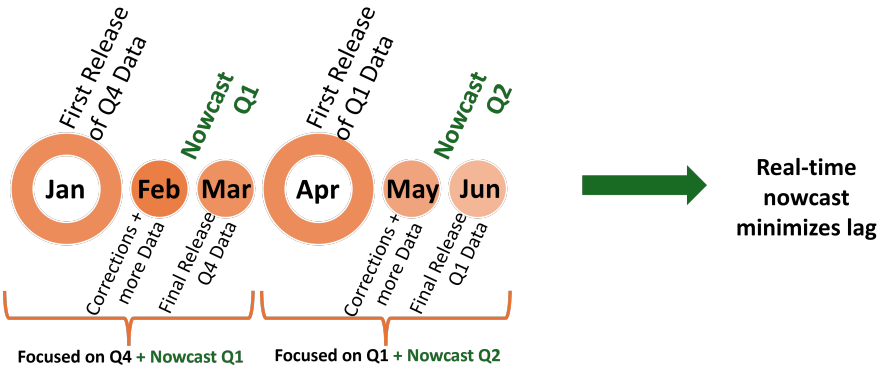


Figure 19: Nowcast timing within a quarter.

This nowcasting layer is implemented jointly with a missing-value imputation step to obtain a balanced real-time panel in the spirit of the Tall-Wide (Bai & Ng, 2021) / Tall-Project (Cahan, Bai, & Ng, 2023) approach.

G.2.2 Mixed-frequency dynamic factor model on the monthly panel

Let $x_{i,t}$ denote the level of series $i = 1, \dots, N_M$ observed at monthly frequency for $t = 1, \dots, T$. Each series is transformed to approximate stationarity using the series-specific operator $T_i(\cdot)$ (indexed by `tcode`, i.e., first differences, log-difference, levels, etc.):

$$y_{i,t} = T_i(x_{i,t}), \qquad i = 1, \dots, N_M, \quad t = 1, \dots, T. \tag{38}$$

Stack $y_t = (y_{1,t}, \dots, y_{N_M,t})'$ to obtain the N_M-dimensional monthly panel. Because the data are observed as real-time vintages (ALFRED / FRED-MD style), the panel is unbalanced and exhibits a ragged edge.

I assume that the transformed monthly panel is driven by a small number of dynamic common factors f_t and idiosyncratic components ε_t:

$$y_t = \Lambda f_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, R), \tag{39}$$

$$f_t = \Phi_1 f_{t-1} + \cdots + \Phi_p f_{t-p} + u_t, \qquad u_t \sim \mathcal{N}(0, Q), \tag{40}$$

where Λ is the $N_M \times r$ loading matrix and p is the number of factor lags. The number of factors r and lags p is selected by information criteria in the spirit of Bai and Ng (2002); Stock and Watson (2002), using the Hannan–Quinn criterion (Hannan & Quinn, 1979).

The model (39)–(40) is estimated in a quasi–maximum likelihood framework using a two-step procedure: principal components provide initial estimates of f_t and Λ, which are then refined by an EM algorithm with Kalman filtering and smoothing (Doz et al., 2011, 2012; Kalman, 1960; Kalman & Bucy, 1961). This approach naturally accommodates missing observations in y_t and delivers imputations as model-implied conditional expectations.

Let $\widehat{f_t}$ denote the estimated monthly factors and $\widehat{y}_{i,t}$ the corresponding fitted values for the transformed series, including imputations for missing entries.

G.2.3 Quarterly aggregation and bridge regressions

Let quarters be indexed by $q = 1, \dots, T_Q$, and let $\mathcal{M}(q)$ denote the set of months belonging to quarter q. Quarterly factors are obtained by averaging the estimated monthly factors:

$$f_q = \frac{1}{|\mathcal{M}(q)|} \sum_{t \in \mathcal{M}(q)} \widehat{f_t}. \tag{41}$$

A fully-fledged mixed-frequency DFM that jointly treats all quarterly variables in the state vector is computationally infeasible in this setting due to the large number of quarterly series. Instead, I transfer the high-frequency information to quarterly variables via factor-augmented bridge regressions.

Let $z_{\ell,q}$, $\ell = 1, \dots, N_Q$, denote quarterly variables in levels, and define transformed quarterly series

$$y^{\mathrm{Q}}_{\ell,q} = T_\ell(z_{\ell,q}). \tag{42}$$

For each ℓ, estimate an OLS bridge regression on the quarterly factors using only quarters with observed data:

$$y^{\mathrm{Q}}_{\ell,q} = \alpha_\ell + \beta_\ell' f_q + e_{\ell,q}, \qquad q \in Q_\ell, \tag{43}$$

where Q_ℓ is the set of quarters for which $y^{\mathrm{Q}}_{\ell,q}$ is observed. For quarters where $y^{\mathrm{Q}}_{\ell,q}$ is missing but factors are available (including the nowcast quarter), form predicted transformed values

$$\widehat{y}^{\mathrm{Q}}_{\ell,q} = \widehat{\alpha}_\ell + \widehat{\beta}'_\ell f_q. \tag{44}$$

These predictions fill gaps in transformed space and are back-transformed to levels via $T_\ell^{-1}(\cdot)$:

$$z^{\mathrm{Q}}_{\ell,q} = \begin{cases} z_{\ell,q}, & \text{if } z_{\ell,q} \text{ is observed,} \\ \widehat{z}_{\ell,q}, & \text{if } z_{\ell,q} \text{ is missing,} \end{cases} \tag{45}$$

where $\widehat{z}_{\ell,q} = T_\ell^{-1}(\widehat{y}^{\mathrm{Q}}_{\ell,q}$; anchored in $z_{\ell,\cdot})$. This yields a completed quarterly panel that incorporates information from monthly indicators through f_q, in the spirit of standard factor-based nowcasting frameworks (Bańbura, Giannone, Modugno, & Reichlin, 2013; Giannone et al., 2008).

G.2.4 Institutional nowcasts and override rule

To incorporate publicly available professional and institutional nowcasts for key variables, I additionally use nowcasts from the Survey of Professional Forecasters (SPF), the Cleveland Fed inflation nowcasting project, and the Atlanta Fed GDPNow. When multiple nowcasts are available for the same variable and vintage, institutional nowcasts are preferred over SPF. When such external nowcasts are available at the vintage date, they *override* the factor-implied nowcasts for the corresponding variable and quarter; otherwise, the DFM/bridge imputations are used.

G.3 The Dataset

In the Tables 13-31, analogue to McCracken and Ng (2021), the series adjusted, all series' descriptions, as well as variables dropping out over time are documented. Red entries mark corrections from the latest online appendix[1] by McCracken and Ng (2021), where , for example, the incorrect deflator was noted in the appendix for generating the latest published data.

[1] As of August 2026

G.3.1 Series adjusted by FRED-QD+

Table 13: Series adjusted by FRED-QD+

ID	Variable	Adjustment
60	Unemployment rate (< 27 weeks)	(UNEMPLOY - UEMP27OV)/CLF16OV
61	Unemployment rate (> 27 weeks)	UEMP27OV/CLF16OV
80	Help-wanted index	Splice LMJVTTUVUSM647S with Barnichon (2010) series
88	Real manu. and trade	(i) Adjust M0602BUSM144NNBR for inflation using PCEPI (ii) Seasonal adjust with ARIMA X12 (iii) Splice with NAICS series CMRMTSPL
89	Retail/food sales	Splice SIC series RETAIL with NAICS series RSAFS
90	New orders (durables)	Splice SIC series AMDMNO and NAICS series DGORDER
92	Unfilled orders (durables)	Splice SIC series AMDMUO and NAICS series AMDMUO
93	New orders (nondefense)	Splice SIC series ANDENO and NAICS series ANDENO
130	Crude oil	Splice OILPRICE with MCOILWTICO
153	30yr mortgage to 10yr Treasury	MRTG - GS10
154	6mth T-bill – 3mth T-bill	TB6MS - TB3MS
155	1yr Treasury – 3mth T-bill	GS1 - TB3MS
156	10yr Treasury – 3mth T-bill	GS10 - TB3MS
157	3mth commercial – 3mth T-bill	CPF3M - TB3MS
172	Household/nonprof liab to income	TLBSHNO/DSPI
174	Household/nonprof net worth to income	TNWBSHNO/DSPI
175	Assets net of Real Estate	(TABSHNO - HNOREMQ027S / 1000)/PCEPILFE
178	S&P 100 volatility: VXO	Splice Bloom (2009) series with VXOCLS
184	Switzerland/U.S. FX	Filled back to 1959 from Banking/Monetary statistics
185	Japan/U.S. FX	Filled back to 1959 from Banking/Monetary statistics
186	U.K./U.S. FX	Filled back to 1959 from Banking/Monetary statistics
187	Cdn/U.S. FX	Filled back to 1959 from Banking/Monetary statistics
188	Consumer sentiment	Splice UMSCENT1 with UMSCENT
220	Help wanted to unemployed	HWI/UNEMPLOY
221	Initial claims	Splice monthly series M08297USM548NNBR with weekly ICSA
222	Business inventories	Splice SIC series and NAICS series BUSINV
223	Inventory to sales	Splice SIC series and NAICS series ISRATIO
224	Consumer credit to P.I.	NONREVSL/PI
235	Business liabilities to income	TLBSNNCB/(CNCF - A054RC1Q027SBEA)
238	Business net worth to income	TNWMVBSNNCB/(CNCF - A054RC1Q027SBEA)
240	NonCorp busi. liabilities to income	TLBSNNB/(CNCF - A054RC1Q027SBEA)
243	NonCorp busi. net worth to income	TNWBSNNB/(CNCF - A054RC1Q027SBEA)
244	Business income	(CNCF - A054RC1Q027SBEA)/IPDBS
249	Baa-Aaa Spread	BAA-AAA
250	Baa-FFM Spread	BAA-FEDFUNDS
251	10-Year-2-Year Treasury Yield Spread	GS10-GS2
252	1-Year Treasury Federal Funds Spread	GS1-FEDFUNDS
253	Aaa-10 Year Treasury Yield Spread	AAA-GS10
258	Output Gap	(GDPC1-GDPPOT)/GDPPOT
259	Barro Misery Index	ΔFEDFUNDS+ΔUNRATE+ΔCPIAUCSL-GAPx

G.3.2 Variables used in FRED-QD+

Table 14: Group 1: NIPA

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
1	1	5	0	GDPC1	GDP	Real Gross Domestic Product, 3 Decimal (Billions of Chained 2017 Dollars)
2	2	5	0	PCECC96	Consumption	Real Personal Consumption Expenditures (Billions of Chained 2017 Dollars)
3	3	5	1	PCDGx	Cons:Dur	Real personal consumption expenditures: Durable goods (Billions of Chained 2017 Dollars), deflated using its own price index
4	4	5	1	PCESVx	Cons:Svc	Real Personal Consumption Expenditures: Services (Billions of 2017 Dollars), deflated using its own price index
5	5	5	1	PCNDx	Cons:NonDur	Real Personal Consumption Expenditures: Nondurable Goods (Billions of Chained 2017 Dollars), deflated using its own price index
6	6	5	0	GPDIC1	Investment	Real Gross Private Domestic Investment, 3 decimal (Billions of Chained 2017 Dollars)
7	7	5	0	FPIx	FixedInv	Real private fixed investment (Billions of Chained 2017 Dollars), deflated using its own price index
8	8	5	1	Y033RC1Q027SBEAx	Inv:Equip&Software	Real Gross Private Domestic Investment: Fixed Investment: Nonresidential: Equipment (Billions of Chained 2017 Dollars), deflated using its own price index
9	9	5	1	PNFIx	FixInv:NonRes	Real private fixed investment: Nonresidential (Billions of Chained 2017 Dollars), deflated using its own price index
10	10	5	1	PRFIx	FixedInv:Res	Real private fixed investment: Residential (Billions of Chained 2017 Dollars), deflated using its own price index
11	11	1	1	A014RE1Q156NBEA	Inv:Inventories	Shares of gross domestic product: Gross private domestic investment: Change in private inventories (Percent)
12	12	5	0	GCEC1	Gov.Spending	Real Government Consumption Expenditures & Gross Investment (Billions of Chained 2017 Dollars)
13	13	1	1	A823RL1Q225SBEA	Gov:Fed	Real Government Consumption Expenditures and Gross Investment: Federal (Percent Change from Preceding Period)
14	14	5	1	FGRECPTx	Real Gov Receipts	Real Federal Government Current Receipts (Billions of Chained 2017 Dollars), deflated using GDP Deflator
15	15	5	1	SLCEx	Gov:State&Local	Real government state and local consumption expenditures (Billions of Chained 2017 Dollars), deflated using its own price index
16	16	5	1	EXPGSC1	Exports	Real Exports of Goods & Services, 3 Decimal (Billions of Chained 2017 Dollars)
17	17	5	1	IMPGSC1	Imports	Real Imports of Goods & Services, 3 Decimal (Billions of Chained 2017 Dollars)
18	18	5	0	DPIC96	Disp-Income	Real Disposable Personal Income (Billions of Chained 2017 Dollars)
19	19	5	0	OUTNFB	Output:NFB	Nonfarm Business Sector: Real Output (Index 2017=100)
20	20	5	0	OUTBS	Output:Bus	Business Sector: Real Output (Index 2017=100)
21	21	5	0	OUTMS	Output:Manuf	Manufacturing Sector: Real Output (Index 2017=100)
192	n.a.	2	0	GFDEGDQ188S		Federal Debt: Total Public Debt as Percent of GDP (Percent)
193	n.a.	2	0	GFDEBTNx		Real Federal Debt: Total Public Debt (Millions of 2017 Dollars), deflated by PCE

Table 15: Group 2: Industrial Production

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
22	22	5	0	INDPRO	IP:Total index	Industrial Production Index (Index 2017=100)
23	23	5	0	IPFINAL	IP:Final products	Industrial Production: Final Products (Market Group) (Index 2017=100)
24	24	5	0	IPCONGD	IP:Consumer goods	Industrial Production: Consumer Goods (Index 2017=100)
25	25	5	0	IPMAT	IP:Materials	Industrial Production: Materials (Index 2017=100)
26	26	5	1	IPDMAT	IP:Dur gds materials	Industrial Production: Durable Materials (Index 2017=100)
27	27	5	1	IPNMAT	IP:Nondur gds materials	Industrial Production: Nondurable Materials (Index 2017=100)
28	28	5	1	IPDCONGD	IP:Dur Cons. Goods	Industrial Production: Durable Consumer Goods (Index 2017=100)
29	29	5	1	IPB51110SQ	IP:Auto	Industrial Production: Durable Goods: Automotive products (Index 2017=100)
30	30	5	1	IPNCONGD	IP:NonDur Cons Goods	Industrial Production: Nondurable Consumer Goods (Index 2017=100)
31	31	5	1	IPBUSEQ	IP:Bus Equip	Industrial Production: Business Equipment (Index 2017=100)
32	32	5	1	IPB51220SQ	IP:Energy Prds	Industrial Production: Consumer energy products (Index 2017=100)
33	33	1	1	TCU	Capu Tot	Capacity Utilization: Total Industry (Percent of Capacity)
34	34	1	1	CUMFNS	Capu Man.	Capacity Utilization: Manufacturing (SIC) (Percent of Capacity)
194	n.a.	5	0	IPMANSICS		Industrial Production: Manufacturing (SIC) (Index 2017=100)
195	n.a.	5	0	IPB51222S		Industrial Production: Residential Utilities (Index 2017=100)
196	n.a.	5	0	IPFUELS		Industrial Production: Fuels (Index 2017=100)

Table 16: Group 3: Employment and Unemployment

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
35	35	5	0	PAYEMS	Emp:Nonfarm	All Employees: Total nonfarm (Thousands of Persons)
36	36	5	0	USPRIV	Emp:Private	All Employees: Total Private Industries (Thousands of Persons)
37	37	5	0	MANEMP	Emp:mfg	All Employees: Manufacturing (Thousands of Persons)
38	38	5	0	SRVPRD	Emp:Services	All Employees: Service-Providing Industries (Thousands of Persons)
39	39	5	0	USGOOD	Emp:Goods	All Employees: Goods-Producing Industries (Thousands of Persons)
40	40	5	1	DMANEMP	Emp:DurGoods	All Employees: Durable goods (Thousands of Persons)
41	41	5	0	NDMANEMP	Emp:Nondur Goods	All Employees: Nondurable goods (Thousands of Persons)
42	42	5	1	USCONS	Emp:Const	All Employees: Construction (Thousands of Persons)
43	43	5	1	USEHS	Emp:Edu&Health	All Employees: Education & Health Services (Thousands of Persons)
44	44	5	1	USFIRE	Emp:Finance	All Employees: Financial Activities (Thousands of Persons)
45	45	5	1	USINFO	Emp:Infor	All Employees: Information Services (Thousands of Persons)
46	46	5	1	USPBS	Emp:Bus Serv	All Employees: Professional & Business Services (Thousands of Persons)
47	47	5	1	USLAH	Emp:Leisure	All Employees: Leisure & Hospitality (Thousands of Persons)
48	48	5	1	USSERV	Emp:OtherSvcs	All Employees: Other Services (Thousands of Persons)
49	49	5	1	USMINE	Emp:Mining/NatRes	All Employees: Mining and logging (Thousands of Persons)
50	50	5	1	USTPU	Emp:Trade&Trans	All Employees: Trade, Transportation & Utilities (Thousands of Persons)
51	51	5	0	USGOVT	Emp:Gov	All Employees: Government (Thousands of Persons)
52	52	5	1	USTRADE	Emp:Retail	All Employees: Retail Trade (Thousands of Persons)
53	53	5	1	USWTRADE	Emp:Wholesal	All Employees: Wholesale Trade (Thousands of Persons)
54	54	5	1	CES9091000001	Emp:Gov(Fed)	All Employees: Government: Federal (Thousands of Persons)
55	55	5	1	CES9092000001	Emp:Gov (State)	All Employees: Government: State Government (Thousands of Persons)
56	56	5	1	CES9093000001	Emp:Gov (Local)	All Employees: Government: Local Government (Thousands of Persons)
57	57	5	0	CE16OV	Emp:Total (HHSurve)	Civilian Employment (Thousands of Persons)
58	58	2	0	CIVPART	LF Part Rate	Civilian Labor Force Participation Rate (Percent)
59	59	2	0	UNRATE	Unemp Rate	Civilian Unemployment Rate (Percent)
60	60	2	0	UNRATESTx	Urate_ST	Unemployment Rate less than 27 weeks (Percent)
61	61	2	0	UNRATELTx	Urate_LT	Unemployment Rate for more than 27 weeks (Percent)
62	62	2	1	LNS14000012	Urate:Age16-19	Unemployment Rate - 16 to 19 years (Percent)
63	63	2	1	LNS14000025	Urate:Age>20 Men	Unemployment Rate - 20 years and over, Men (Percent)
64	64	2	1	LNS14000026	Urate:Age>20 Women	Unemployment Rate - 20 years and over, Women (Percent)
65	65	5	1	UEMPLT5	U:Dur<5wks	Number of Civilians Unemployed - Less Than 5 Weeks (Thousands of Persons)
66	66	5	1	UEMP5TO14	U:Dur5-14wks	Number of Civilians Unemployed for 5 to 14 Weeks (Thousands of Persons)
67	67	5	1	UEMP15T26	U:dur>15-26wks	Number of Civilians Unemployed for 15 to 26 Weeks (Thousands of Persons)
68	68	5	1	UEMP27OV	U:Dur>27wks	Number of Civilians Unemployed for 27 Weeks and Over (Thousands of Persons)
69	69	5	1	LNS13023621	U:Job losers	Unemployment Level - Job Losers (Thousands of Persons)
70	70	5	1	LNS13023557	U:LF Reenty	Unemployment Level - Reentrants to Labor Force (Thousands of Persons)
71	71	5	1	LNS13023705	U:Job Leavers	Unemployment Level - Job Leavers (Thousands of Persons)
72	72	5	1	LNS13023569	U:New Entrants	Unemployment Level - New Entrants (Thousands of Persons)

Table 17: Group 3: Employment and Unemployment (continued)

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
73	73	5	1	LNS12032194	Emp:SlackWk	Employment Level - Part-Time for Economic Reasons, All Industries (Thousands of Persons)
74	74	5	0	HOABS	EmpHrs:Bus Sec	Business Sector: Hours of All Persons (Index 2017=100)
75	75	5	0	HOAMS	EmpHrs:mfg	Manufacturing Sector: Hours of All Persons (Index 2017=100)
76	76	5	0	HOANBS	EmpHrs:nfb	Nonfarm Business Sector: Hours of All Persons (Index 2017=100)
77	77	1	1	AWHMAN	AWH Man	Average Weekly Hours of Production and Nonsupervisory Employees: Manufacturing (Hours)
78	78	2	1	AWHNONAG	AWH Privat	Average Weekly Hours of Production and Nonsupervisory Employees: Total private (Hours)
79	79	2	1	AWOTMAN	AWH Overtime	Average Weekly Overtime Hours of Production and Nonsupervisory Employees: Manufacturing (Hours)
80	80	1	0	HWIx	HelpWnted	Help-Wanted Index
197	n.a.	2	0	UEMPMEAN		Average (Mean) Duration of Unemployment (Weeks)
198	n.a.	2	0	CES0600000007		Average Weekly Hours of Production and Nonsupervisory Employees: Goods-Producing
220	n.a.	2	0	HWIURATIOx		Ratio of Help Wanted/No. Unemployed
221	n.a.	5	0	CLAIMSx		Initial Claims
261	n.a.	5	0	CNP16OV		Civilian Noninstitutional Population, >16 Years

Table 18: Group 4: Housing

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
81	81	5	0	HOUST	Hstarts	Housing Starts: Total: New Privately Owned Housing Units Started (Thousands of Units)
82	82	5	0	HOUST5F	Hstarts >5units	Privately Owned Housing Starts: 5-Unit Structures or More (Thousands of Units)
83	83	5	1	PERMIT	Hpermits	New Private Housing Units Authorized by Building Permits (Thousands of Units)
84	84	5	1	HOUSTMW	Hstarts:MW	Housing Starts in Midwest Census Region (Thousands of Units)
85	85	5	1	HOUSTNE	Hstarts:NE	Housing Starts in Northeast Census Region (Thousands of Units)
86	86	5	1	HOUSTS	Hstarts:S	Housing Starts in South Census Region (Thousands of Units)
87	87	5	1	HOUSTW	Hstarts:W	Housing Starts in West Census Region (Thousands of Units)
179	190	5	1	USSTHPI	Real Hprice:OFHEO	Real All-Transactions House Price Index for the United States (Index 1980 Q1=100), deflated by Core PCE
180	191	5	1	SPCS10RSA	Real CS_10	S&P/Case-Shiller 10-City Composite Home Price Index (Index January 2000 = 100), deflated by Core PCE
181	192	5	1	SPCS20RSA	Real CS_20	S&P/Case-Shiller 20-City Composite Home Price Index (Index January 2000 = 100), deflated by Core PCE
227	n.a.	5	0	PERMITNE		New Private Housing Units Authorized by Building Permits in the Northeast Census Region (Thousands, SAAR)
228	n.a.	5	0	PERMITMW		New Private Housing Units Authorized by Building Permits in the Midwest Census Region (Thousands, SAAR)
229	n.a.	5	0	PERMITS		New Private Housing Units Authorized by Building Permits in the South Census Region (Thousands, SAAR)
230	n.a.	5	0	PERMITW		New Private Housing Units Authorized by Building Permits in the West Census Region (Thousands, SAAR)

Table 19: Group 5: Inventories, Orders, and Sales

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
88	89	5	0	CMRMTSPLx	MT Sales	Real Manufacturing and Trade Industries Sales (Millions of Chained 2017 Dollars)
89	90	5	1	RSAFSx	Ret. Sale	~~Real~~ Retail and Food Services Sales ~~(Millions of Chained 2017 Dollars), deflated by Core PCE~~
90	91	5	1	AMDMNOx	Orders (DurMfg)	Real Manufacturers' New Orders: Durable Goods (Millions of 2017 Dollars), deflated by Core PCE
91	92	5	1	ACOGNOx	Orders(ConsumerGoods/Mat.)	Real Value of Manufacturers' New Orders for Consumer Goods Industries (Millions of 2017 Dollars), deflated by Core PCE
92	93	5	1	AMDMUOx	UnfOrders(DurGds)	Real Value of Manufacturers' Unfilled Orders for Durable Goods Industries (Million of 2017 Dollars), deflated by Core PCE
93	94	5	1	ANDENOx	Orders(NonDefCap)	Real Value of Manufacturers' New Orders for Capital Goods: Nondefense Capital Goods Industries (Million of 2017 Dollars), deflated by Core PCE
94	96	5	1	INVCQRMTSPL	MT Invent	Real Manufacturing and Trade Inventories (Millions of 2017 Dollars)
222	n.a.	5	0	BUSINVx		Total Business Inventories (Millions of Dollars)
223	n.a.	2	0	ISRATIOx		Total Business: Inventories to Sales Ratio

Table 20: Group 6: Prices

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
95	97	6	0	PCECTPI	PCED	Personal Consumption Expenditures: Chain-type Price Index (Index 2017=100)
96	98	6	0	PCEPILFE	PCED_LFE	Personal Consumption Expenditures Excluding Food and Energy (Chain-Type Price Index) (Index 2017=100)
97	99	6	0	GDPCTPI	GDP Defl	Gross Domestic Product: Chain-type Price Index (Index 2017=100)
98	100	6	1	GPDICTPI	GPDI Defl	Gross Private Domestic Investment: Chain-type Price Index (Index 2017=100)
99	101	6	1	IPDBS	BusSec Defl	Business Sector: Implicit Price Deflator (Index 2017=100)
100	102	6	0	DGDSRG3Q086SBEA	PCED_Goods	Personal consumption expenditures: Goods (chain-type price index)
101	103	6	0	DDURRG3Q086SBEA	PCED_DurGoods	Personal consumption expenditures: Durable goods (chain-type price index)
102	104	6	0	DSERRG3Q086SBEA	PCED_Serv	Personal consumption expenditures: Services (chain-type price index)
103	105	6	0	DNDGRG3Q086SBEA	PCED_NDurGoods	Personal consumption expenditures: Nondurable goods (chain-type price index)
104	106	6	0	DHCERG3Q086SBEA	PCED_HouseholdServ.	Personal consumption expenditures: Services: Household consumption expenditures (chain-type price index)
105	107	6	1	DMOTRG3Q086SBEA	PCED_MotorVec	Personal consumption expenditures: Durable goods: Motor vehicles and parts (chain-type price index)
106	108	6	1	DFDHRG3Q086SBEA	PCED_DurHousehold	Personal consumption expenditures: Durable goods: Furnishings and durable household equipment (chain-type price index)
107	109	6	1	DREQRG3Q086SBEA	PCED_Recreation	Personal consumption expenditures: Durable goods: Recreational goods and vehicles (chain-type price index)
108	110	6	1	DODGRG3Q086SBEA	PCED_OthDurGds	Personal consumption expenditures: Durable goods: Other durable goods (chain-type price index)
109	111	6	1	DFXARG3Q086SBEA	PCED_Food_Bev	Personal consumption expenditures: Nondurable goods: Food and beverages purchased for off-premises consumption (chain-type price index)
110	112	6	1	DCLORG3Q086SBEA	PCED_Clothing	Personal consumption expenditures: Nondurable goods: Clothing and footwear (chain-type price index)
111	113	6	1	DGOERG3Q086SBEA	PCED_Gas_Enrgy	Personal consumption expenditures: Nondurable goods: Gasoline and other energy goods (chain-type price index)
112	114	6	1	DONGRG3Q086SBEA	PCED_OthNDurGds	Personal consumption expenditures: Nondurable goods: Other nondurable goods (chain-type price index)
113	115	6	1	DHUTRG3Q086SBEA	PCED_Housing-Utilities	Personal consumption expenditures: Services: Housing and utilities (chain-type price index)
114	116	6	1	DHLCRG3Q086SBEA	PCED_HealthCare	Personal consumption expenditures: Services: Health care (chain-type price index)
115	117	6	1	DTRSRG3Q086SBEA	PCED_TransSvg	Personal consumption expenditures: Transportation services (chain-type price index)
116	118	6	1	DRCARG3Q086SBEA	PCED_RecServices	Personal consumption expenditures: Recreation services (chain-type price index)
117	119	6	1	DFSARG3Q086SBEA	PCED_FoodServ_Acc.	Personal consumption expenditures: Services: Food services and accommodations (chain-type price index)
118	120	6	1	DIFSRG3Q086SBEA	PCED_FIRE	Personal consumption expenditures: Financial services and insurance (chain-type price index)
119	121	6	1	DOTSRG3Q086SBEA	PCED_OtherServices	Personal consumption expenditures: Other services (chain-type price index)
120	122	6	0	CPIAUCSL	CPI	Consumer Price Index for All Urban Consumers: All Items (Index 1982–84=100)
121	132	6	0	CPILFESL	CPI_LFE	Consumer Price Index for All Urban Consumers: All Items Less Food & Energy (Index 1982–84=100)
122	124	6	0	WPSFD49207	PPI:FinGds	Producer Price Index by Commodity for Finished Goods (Index 1982=100)
132	125	6	0	PPIACO	PPI	Producer Price Index for All Commodities (Index 1982=100)
124	132	6	1	WPSFD49502	PPI:FinConsGds	Producer Price Index: Finished Consumer Goods (Index 1982=100)
125	127	6	1	WPSFD4111	PPI:FinConsGds(Food)	Producer Price Index: Finished Consumer Foods (Index 1982=100)
132	128	6	1	PPIIDC	PPI:IndCom	Producer Price Index: Industrial Commodities (Index 1982=100)
127	129	6	1	WPSID61	PPI:IntMat	Producer Price Index: Intermediate Materials, Supplies & Components (Index 1982=100)
128	133	5	1	WPU0531	Real Price:NatGas	Producer Price Index: Natural Gas (Index 1982=100), deflated by Core PCE
129	134	5	1	WPU0561	Real Price:Oil	Producer Price Index: Crude Petroleum (Domestic Production) (Index 1982=100), deflated by Core PCE
130	135	5	0	OILPRICEx	Real Crude Oil Price	Real Crude Oil Prices: WTI (2017 Dollars per Barrel), deflated by Core PCE
205	n.a.	6	0	WPSID62		Producer Price Index: Crude Materials for Further Processing (Index 1982=100)
206	n.a.	6	0	PPICMM		Producer Price Index: Metals and primary nonferrous metals (Index 1982=100)
207	n.a.	6	0	CPIAPPSL		Consumer Price Index: Apparel (Index 1982–84=100)
208	n.a.	6	0	CPITRNSL		Consumer Price Index: Transportation (Index 1982–84=100)
209	n.a.	6	0	CPIMEDSL		Consumer Price Index: Medical Care (Index 1982–84=100)
210	n.a.	6	0	CUSR0000SAC		Consumer Price Index: Commodities (Index 1982–84=100)
260	n.a.	5	0	MICH		Median expected price change next 12 months, Survey of Consumers, University of Michigan

Table 21: Group 6: Prices (continued)

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
211	n.a.	6	0	CUSR0000SAD		Consumer Price Index for All Urban Consumers: Durables (Index 1982–84=100)
212	n.a.	6	0	CUSR0000SAS		Consumer Price Index for All Urban Consumers: Services (Index 1982–84=100)
213	n.a.	6	0	CPIULFSL		Consumer Price Index for All Urban Consumers: All Items Less Food (Index 1982–84=100)
214	n.a.	6	0	CUSR0000SA0L2		Consumer Price Index for All Urban Consumers: All items less shelter (Index 1982–84=100)
215	n.a.	6	0	CUSR0000SA0L5		Consumer Price Index for All Urban Consumers: All items less medical care (Index 1982–84=100)
233	n.a.	6	0	CUSR0000SEHC		CPI for All Urban Consumers: Owners' equivalent rent of residences (Index Dec 1982=100)

Table 22: Group 7: Earnings and Productivity

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
131	136	5	0	AHETPIx	Real AHE:PrivInd	Real Avg Hourly Earnings: Production & Nonsupervisory Employees, Total Private (2017 Dollars), deflated by Core PCE
132	137	5	0	CES2000000008x	Real AHE:Const	Real Avg Hourly Earnings: Construction (2017 Dollars), deflated by Core PCE
133	138	5	0	CES3000000008x	Real AHE:MFG	Real Avg Hourly Earnings: Manufacturing (2017 Dollars), deflated by Core PCE
134	139	5	1	COMPRMS	CPH:Mfg	Manufacturing Sector: Real Compensation Per Hour (Index 2017=100)
135	140	5	1	COMPRNFB	CPH:NFB	Nonfarm Business Sector: Real Compensation Per Hour (Index 2017=100)
136	141	5	1	RCPHBS	CPH:Bus	Business Sector: Real Compensation Per Hour (Index 2017=100)
137	142	5	1	OPHMFG	OPH:mfg	Manufacturing Sector: Real Output Per Hour of All Persons (Index 2017=100)
138	143	5	1	OPHNFB	OPH:nfb	Nonfarm Business Sector: Real Output Per Hour of All Persons (Index 2017=100)
139	144	5	0	OPHPBS	OPH:Bus	Business Sector: Real Output Per Hour of All Persons (Index 2017=100)
140	145	5	0	ULCBS	ULC:Bus	Business Sector: Unit Labor Cost (Index 2017=100)
141	146	5	1	ULCMFG	ULC:Mfg	Manufacturing Sector: Unit Labor Cost (Index 2017=100)
142	147	5	1	ULCNFB	ULC:NFB	Nonfarm Business Sector: Unit Labor Cost (Index 2017=100)
143	148	5	1	UNLPNBS	UNLPay:nfb	Nonfarm Business Sector: Unit Nonlabor Payments (Index 2017=100)
216	n.a.	6	0	CES0600000008		Avg Hourly Earnings of Production & Nonsupervisory Employees: Goods-Producing (Dollars per Hour)

Table 23: Group 8: Interest Rates

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
144	149	2	1	FEDFUNDS	FedFunds	Effective Federal Funds Rate (Percent)
145	150	2	1	TB3MS	TB-3Mth	3-Month Treasury Bill: Secondary Market Rate (Percent)
146	151	2	0	TB6MS	TM-6MTH	6-Month Treasury Bill: Secondary Market Rate (Percent)
147	153	2	0	GS1	TB-1YR	1-Year Treasury Constant Maturity Rate (Percent)
148	154	2	0	GS10	TB-10YR	10-Year Treasury Constant Maturity Rate (Percent)
149	155	2	0	MORTGAGE30US	Mort-30Yr	30-Year Conventional Mortgage Rate© (Percent)
150	156	2	0	AAA	AAA Bond	Moody's Seasoned Aaa Corporate Bond Yield© (Percent)
151	157	2	0	BAA	BAA Bond	Moody's Seasoned Baa Corporate Bond Yield© (Percent)
152	158	1	1	BAA10YM	BAA_GS10	Baa Corporate Bond Yield Relative to 10-Year Treasury (Percent)
153	159	1	1	MORTG10YRx	MRTG_GS10	30-Year Mortgage Rate Relative to 10-Year Treasury (Percent)
154	160	1	1	TB6M3Mx	tb6m_tb3m	6-Month Treasury Bill Minus 3-Month Treasury Bill (Percent)
155	161	1	1	GS1TB3Mx	GS1_tb3m	1-Year Treasury Minus 3-Month Treasury Bill (Percent)
156	162	1	1	GS10TB3Mx	GS10_tb3m	10-Year Treasury Minus 3-Month Treasury Bill (Percent)
157	163	1	1	CPF3MTB3Mx	CP_Tbill Spread	3-Month Commercial Paper Minus 3-Month T-bill (Percent)
201	n.a.	2	0	GS5		5-Year Treasury Constant Maturity Rate (Percent)
202	n.a.	1	0	TB3SMFFM		3-Month Treasury Constant Maturity Minus Federal Funds Rate (Percent)
203	n.a.	1	0	T5YFFM		5-Year Treasury Constant Maturity Minus Federal Funds Rate (Percent)
204	n.a.	1	0	AAAFFM		Aaa Corporate Bond Yield Minus Federal Funds Rate (Percent)
225	n.a.	2	0	CP3M		3-Month AA Financial Commercial Paper Rate (Percent)
226	n.a.	1	0	COMPAPFF		3-Month Commercial Paper Minus Federal Funds Rate (Percent)
249	n.a.	1	0	BAAAAAx		Moody's Seasoned Baa Corporate Bond Yield minus Moody's Seasoned Aaa Corporate Bond Yield
250	n.a.	1	0	BAAFFM		Moody's Seasoned Baa Corporate Bond Minus Federal Funds Rate
251	n.a.	1	0	T10Y2Yx		10-Year Treasury Constant Maturity Minus 2-Year Treasury Constant Maturity
252	n.a.	1	0	T1YFFM		1-Year Treasury Constant Maturity Minus Federal Funds Rate
253	n.a.	1	0	AAA10YMx		Moody's Seasoned Aaa Corporate Bond Minus 10-Year Treasury Constant Maturity

Table 24: Group 9: Money and Credit

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
158	167	5	0	BOGMBASEREALx	Real Mbase	St. Louis Adjusted Monetary Base (Billions of 1982–84 Dollars), deflated by CPI
160	169	5	0	M1REAL	Real m1	Real M1 Money Stock (Billions of 1982–84 Dollars), deflated by CPI
161	170	5	0	M2REAL	Real m2	Real M2 Money Stock (Billions of 1982–84 Dollars), deflated by CPI
163	172	5	1	BUSLOANSx	Real C&Lloand	Real Commercial & Industrial Loans, All Commercial Banks (Billions of 2017 Dollars), deflated by Core PCE
164	173	5	1	CONSUMERx	Real ConsLoans	Real Consumer Loans at All Commercial Banks (Billions of 2017 Dollars), deflated by Core PCE
165	174	5	1	NONREVSLx	Real NonRevCredit	Total Real Nonrevolving Credit Outstanding (Billions of Dollars), deflated by Core PCE
166	175	5	1	REALLNx	Real LoansRealEst	Real Estate Loans, All Commercial Banks (Billions of 2017 Dollars), deflated by Core PCE
167	176	5	1	REVOLSLx	Real RevolvCredit	Real Revolving Credit Outstanding (Billions of 2017 Dollars), deflated by Core PCE
168	177	5	0	TOTALSLx	Real ConsuCred	Total Consumer Credit Outstanding, deflated by Core PCE
169	178	1	1	DRIWCIL	FRBSLO_Consumers	FRB Senior Loan Officer Survey: Net % Reporting Increased Willingness to Make Consumer Installment Loans
199	n.a.	6	0	TOTRESNS		Total Reserves of Depository Institutions (Billions of Dollars)
200	n.a.	7	0	NONBORRES		Nonborrowed Reserves of Depository Institutions (Millions of Dollars)
217	n.a.	6	0	DTCOLNVHFNM		Consumer Motor Vehicle Loans Owned by Finance Companies (Millions of Dollars)
218	n.a.	6	0	DTCTHFNM		Total Consumer Loans & Leases Owned and Securitized by Finance Companies (Millions of Dollars)
219	n.a.	6	0	INVEST		Securities in Bank Credit at All Commercial Banks (Billions of Dollars)
254	n.a.	2	0	DRTSCILM		Net Percentage of Domestic Banks Tightening Standards for Commercial and Industrial Loans to Large and Middle-Market Firms
255	n.a.	2	0	DRTSCIS		Net Percentage of Domestic Banks Tightening Standards for Commercial and Industrial Loans to Small Firms
256	n.a.	2	0	DRTSCFLM		Net Percentage of Domestic Banks Increasing Spreads of Loan Rates Over Banks' Cost of Funds to Large and Middle-Market Firms

Table 25: Group 10: Household Balance Sheets

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
170	179	5	0	TABSHNOx	Real HHW:TASA	Real Total Assets of Households & Nonprofit Organizations (Billions of 2017 Dollars), deflated by Core PCE
171	181	5	1	TLBSHNOx	Real HHW:LiabSA	Real Total Liabilities of Households & Nonprofit Organizations (Billions of 2017 Dollars), deflated by Core PCE
172	182	5	0	LIABPIx	liab_PDISA	Liabilities of Households & Nonprofit Orgs Relative to Disposable Personal Income (Percent)
173	183	5	1	TNWBSHNOx	Real HHW:WSA	Real Net Worth of Households & Nonprofit Organizations (Billions of 2017 Dollars), deflated by Core PCE
174	184	1	0	NWPIx	W_PDISA	Net Worth of Households & Nonprofit Organizations Relative to Disposable Personal Income (Percent)
175	185	5	1	TARESAx	Real HHW:TA_RESA	Real Assets of Households & Nonprofit Orgs excl. Real Estate Assets (Billions of 2017 Dollars), deflated by Core PCE
176	186	5	1	HNOREMQ027Sx	Real HHW:RESA	Real Real Estate Assets of Households & Nonprofit Organizations (Billions of 2017 Dollars), deflated by Core PCE
177	188	5	1	TFAABSHNOx	Real HHW:FinSA	Real Total Financial Assets of Households & Nonprofit Organizations (Billions of 2017 Dollars), deflated by Core PCE
224	n.a.	2	0	CONSPIx		Nonrevolving Consumer Credit Relative to Personal Income
262	n.a.	2	0	W875RX1		Real personal income excluding current transfer receipts (Billions of 2017 Dollars)

Table 26: Group 11: Exchange Rates

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
182	193	5	1	TWEXAFEGSMTHx	Ex rate:major	Trade Weighted U.S. Dollar Index: Major Currencies (Index Mar 1973=100)
183	194	5	1	EXUSEU	Ex rate:Euro	U.S. / Euro Foreign Exchange Rate (USD per Euro)
184	195	5	1	EXSZUSx	Ex rate:Switz	Switzerland / U.S. Foreign Exchange Rate
185	196	5	1	EXJPUSx	Ex rate:Japan	Japan / U.S. Foreign Exchange Rate
186	197	5	1	EXUSUKx	Ex rate:UK	U.S. / U.K. Foreign Exchange Rate
187	198	5	1	EXCAUSx	Ex rate:Canada	Canada / U.S. Foreign Exchange Rate

Table 27: Group 12: Other

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
188	199	1	1	UMCSENTx	Cons. Expectations	University of Michigan: Consumer Sentiment (Index 1966 Q1=100)
189	200	2	1	USEPUINDXM	PolicyUncertainty	Economic Policy Uncertainty Index for the United States
257	n.a.	1	0	SAHMREALTIME		Indicator signals the start of a recession when the three-month moving average of UNRATE rises by 0.50 percentage points or more relative to the minimum of the three-month averages from the previous 12 months.
258	n.a.	1	0	GAPx		Output Gap between Real GDP and Real Potential GDP
259	n.a.	1	0	MISERYx		Barro Misery Index (1999) adding up the 1-year change in GAPx, CPI inflation, UNRATE, and FEDFUNDS

Table 28: Group 13: Stock Markets

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
178	189	1	1	VIXCLSx	VIX	CBOE Volatility Index: VIX
231	n.a.	5	0	NIKKEI225		Nikkei Stock Average
232	n.a.	5	0	NASDAQCOM		NASDAQ Composite (Index Feb 5, 1971=100)
245	180	5	0	S&P 500		S&P's Common Stock Price Index: Composite
247	n.a.	2	0	S&P div yield		S&P Composite Common Stock: Dividend Yield
248	n.a.	5	0	S&P PE ratio		S&P Composite Common Stock: Price-Earnings Ratio

Table 29: Group 14: Non-Household Balance Sheets

ID	SW ID	TCODE	SW Factors	FRED MNEMONIC	SW MNEMONIC	DESCRIPTION
192	n.a.	2	0	GFDEGDQ188S		Federal Debt: Total Public Debt as % of GDP
193	n.a.	2	0	GFDEBTNx		Real Federal Debt: Total Public Debt (Millions of 2017 Dollars), deflated by PCE
234	n.a.	5	0	TLBSNNCBx		Real Nonfinancial Corporate Business Liabilities (Billions of 2017 Dollars), deflated by IPDBS
235	n.a.	1	0	TLBSNNCBBDIx		Nonfinancial Corporate Business Liabilities to Disposable Business Income (%)
236	n.a.	5	0	TTAABSNNCBx		Real Nonfinancial Corporate Business Assets (Billions of 2017 Dollars), deflated by IPDBS
237	n.a.	5	0	TNWMVBSNNCBx		Real Nonfinancial Corporate Business Net Worth (Billions of 2017 Dollars), deflated by IPDBS
238	n.a.	2	0	TNWMVBSNNCBBDIx		Nonfinancial Corporate Business Net Worth to Disposable Business Income (%)
239	n.a.	5	0	TLBSNNBx		Real Nonfinancial Noncorporate Business Liabilities (Billions of 2017 Dollars), deflated by IPDBS
240	n.a.	1	0	TLBSNNBBDIx		Nonfinancial Noncorporate Business Liabilities to Disposable Business Income (%)
241	n.a.	5	0	TABSNNBx		Real Nonfinancial Noncorporate Business Assets (Billions of 2017 Dollars), deflated by IPDBS
242	n.a.	5	0	TNWBSNNBx		Real Nonfinancial Noncorporate Business Net Worth (Billions of 2017 Dollars), deflated by IPDBS
243	n.a.	2	0	TNWBSNNBBDIx		Nonfinancial Noncorporate Business Net Worth to Disposable Business Income (%)
244	n.a.	5	0	CNCFx		Real Disposable Business Income (Corporate cash flow w/ IVA minus taxes), 2017 Dollars, deflated by IPDBS
263	n.a.	5	0	CPROFITx		Corporate Profits with Inventory Valuation Adjustment (IVA) and Capital Consumption Adjustment (CCAdj), 2017 Dollars, deflated by IPDBS

G.3.3 Series dropped and changed during the vintage extension

Table 30: Series Dropping Out or Changed due to no Vintages available

ID	Code	Time	Action	Adjustment
3	PCDGx	Prior to Jan 2013	Deflator changed	Deflated using PCEPI instead
4	PCESVx	Prior to Jan 2013	Deflator changed	Deflated using PCEPI instead
5	PCNDx	Prior to Jan 2013	Deflator changed	Deflated using PCEPI instead
7	FPIx	Prior to Jan 2013	Deflator changed	Deflated using PCEPI instead
8	Y033RC1Q027SBEAx	Prior to Sep 2014	Dropped	-
9	PNFIx	Prior to Jan 2013	Deflator changed	Deflated using PCEPI instead
10	PRFIx	Prior to Jan 2013	Deflator changed	Deflated using PCEPI instead
11	A014RE1Q156NBEA	Prior to Feb 2013	Calculation by hand	CBI/GDP
13	A823RL1Q225SBEA	Prior to Sep 2014	Similar Calculation by hand	Δlog((FNDEFX+FDEFX/GDPCTPI))
15	SLCEx	Prior to Feb 2013	Deflator changed	Deflated using GDPCTPI instead
26	IPDMAT	Prior to Dec 2007	Dropped	-
27	IPNMAT	Prior to Dec 2007	Dropped	-
28	IPDCONGD	Prior to Dec 2007	Dropped	-
29	IPB51110SQ	Prior to Feb 2015	Dropped	-
30	IPNCONGD	Prior to Dec 2007	Dropped	-
31	IPBUSEQ	Prior to Dec 2007	Dropped	-
32	IPB51220SQ	Prior to Feb 2015	Dropped	-
43	USEHS	Prior to Jun 2003	Dropped	-
45	USINFO	Prior to Jun 2003	Dropped	-
46	USPBS	Prior to Jun 2003	Dropped	-
47	USLAH	Prior to Jun 2003	Dropped	-
63	LNS14000012	Prior to May 2011	Dropped	-
64	LNS14000025	Prior to May 2011	Dropped	-
65	LNS14000026	Prior to May 2011	Dropped	-
69	LNS13023621	Prior to May 2011	Dropped	-
70	LNS13023557	Prior to May 2011	Dropped	-
71	LNS13023705	Prior to May 2011	Dropped	-
72	LNS13023569	Prior to May 2011	Dropped	-
73	LNS12032194	Prior to May 2011	Dropped	-
80	HWIx	Prior to 2015	Dropped	-
82	HOUST5F	Prior to Nov 2009	Replaced	Replaced by HOUST2F
94	A823RL1Q225SBEA	Prior to Sep 2014	Similar Calculation by hand	BUSINV/PCEPI
100	DGDSRG3Q086SBEA	Prior to Jan 2013	Dropped	-
101	DDURRG3Q086SBEA	Prior to Jan 2013	Dropped	-
102	DSERRG3Q086SBEA	Prior to Jan 2013	Dropped	-
103	DNDGRG3Q086SBEA	Prior to Jan 2013	Dropped	-
104	DHCERG3Q086SBEA	Prior to Jan 2013	Dropped	-
105	DMOTRG3Q086SBEA	Prior to Jan 2013	Dropped	-
106	DFDHRG3Q086SBEA	Prior to Jan 2013	Dropped	-
107	DREQRG3Q086SBEA	Prior to Jan 2013	Dropped	-
108	DODGRG3Q086SBEA	Prior to Jan 2013	Dropped	-
109	DFXARG3Q086SBEA	Prior to Jan 2013	Dropped	-

Table 31: Series Dropping Out or Changed due to no Vintages available, continued

ID	Code	Time	Action	Adjustment
110	DCLORG3Q086SBEA	Prior to Jan 2013	Dropped	-
111	DGOERG3Q086SBEA	Prior to Jan 2013	Dropped	-
112	DONGRG3Q086SBEA	Prior to Jan 2013	Dropped	-
113	DHUTRG3Q086SBEA	Prior to Jan 2013	Dropped	-
114	DHLCRG3Q086SBEA	Prior to Jan 2013	Dropped	-
115	DTRSRG3Q086SBEA	Prior to Jan 2013	Dropped	-
116	DRCARG3Q086SBEA	Prior to Jan 2013	Dropped	-
117	DFSARG3Q086SBEA	Prior to Jan 2013	Dropped	-
118	DIFSRG3Q086SBEA	Prior to Jan 2013	Dropped	-
119	DOTSRG3Q086SBEA	Prior to Jan 2013	Dropped	-
122	WPSFD49207	Prior to Mar 2015	Dropped	-
124	WPSFD49502	Prior to Mar 2015	Dropped	-
125	WPSFD4111	Prior to Mar 2015	Dropped	-
127	WPSID61	Prior to Mar 2015	Dropped	-
170	TABSHNOx	Prior to June 2010	Dropped	-
171	TLBSHNOx	Prior to June 2010	Dropped	-
172	LIABPIx	Prior to June 2010	Dropped	-
173	TNWBSHNOx	Prior to June 2010	Dropped	-
174	NWPIx	Prior to June 2010	Dropped	-
175	TARESAx	Prior to June 2010	Dropped	-
176	HNOREMQ027Sx	Prior to June 2013	Dropped	-
177	TFAABSHNOx	Prior to June 2010	Dropped	-
181	SPCS20RSA	Prior to 2003[1]	Dropped	-
183	EXUSEU	Prior to 2003[1]	Dropped	-
189	USEPUINDXMx	Prior to June 2013	Dropped	-
190	B020RE1Q156NBEA	Prior to Feb 2013	Calculation by hand	EXPGS/GDP
191	B021RE1Q156NBEA	Prior to Feb 2013	Calculation by hand	IMPGS/GDP
192	GFDEGDQ188S	Prior to Oct 2012	Calculation by hand	GFDEBTN/GDP
195	IPB51222S	Prior to Feb 2015	Dropped	-
196	IPFUELS	Prior to Feb 2015	Dropped	-
205	WPSID62	Prior to Mar 2015	Dropped	-
220	HWIURATIOx	Prior to 2015	Dropped	-
233	CUSR0000SEHC	Prior to April 2011	Dropped	-
234	TLBSNNCBx	Prior to June 2010	Dropped	-
235	TLBSNNCBBDIx	Prior to June 2010	Dropped	-
236	TTAABSNNCBx	Prior to June 2010	Dropped	-
237	TNWMVBSNNCBx	Prior to June 2010	Dropped	-
238	TNWMVBSNNCBBDIx	Prior to June 2010	Dropped	-
239	TLBSNNBx	Prior to June 2010	Dropped	-
240	TLBSNNBBDIx	Prior to June 2010	Dropped	-
241	TABSNNBx	Prior to June 2010	Dropped	-
242	TABSNNBBDIx	Prior to June 2010	Dropped	-
243	TNWBSNNBBDIx	Prior to June 2010	Dropped	-
244	CNCFx	Prior to June 2010	Dropped	-

[1] Dropped due to not enough observations.

IMFS WORKING PAPER SERIES

Recent Issues

244/2026	Balint Tatar, Ph.D., Prof. Volker Wieland, Ph.D.	Growth effects of changes in social security contribution rates in Germany
243/2026	Dr. Matthias Rumpf, Prof. Michael Haliassos, Ph.D., Prof. Dr. Tetyana Kosyakova, Prof. Dr. Thomas Otter	From Humans to Algorithms: How Financial Advice Differs Across Professionals, Peers, and LLMs
242/2026	Hendrik Hegemann, Prof. Volker Wieland, Ph.D.	Scenarios Concerning the Possible Consequences of the Iran War for Euro Area Inflation: Update July 2026
241/2026	Julia Pitters, Prof. Dr. Franz Seitz	Beyond the cost debate: A multidimensional approach to quantify the value of cash for society
240/2026	Otmar Issing	Independence of the central bank – Too much power in the hands of unelected technocrats?
239/2026	Jana Röder, Peter Tillmann, Peter Winker, Jinyeong Yun	Safe-Haven Flows into the German Bond Market and the Role of Policy Disagreement
238/2026	Veronika Grimm, Justus Haucap, Stefan Kolev, Volker Wieland	Eskalation im Nahen Osten: Auswirkungen auf Europa – und was (nicht) zu tun ist
237/2026	Veronika Grimm, Justus Haucap, Stefan Kolev, Volker Wieland	Eine Wachstumsagenda für Deutschland
236/2026	Veronika Grimm, Justus Haucap, Stefan Kolev, Volker Wieland	Die gesetzliche Rentenversicherung: Diagnose und Therapie im Jahr 2025
235/2026	Hendrik Hegemann, Prof. Volker Wieland, Ph.D.	Price Stability & Risks - Benign outlook vs potential concerns
234/2026	Moritz Grebe, M.Sc., Prof. Dr. Peter Tillmann	Inflation Expectations and a Conservative Central Banker: Evidence from a Natural Experiment
233/2026	Hendrik Hegemann, Prof. Volker Wieland, Ph.D.	Scenarios Concerning the Possible Consequences of the Iran War for Euro Area Inflation
232/2026	Michal Grabowski, Iulia Costea	Selected European Law Challenges Related to the Use of Artificial Intelligence Payment Agents
231/2026	Hendrik Hegemann	Energy Price Shocks and Inflation in the Euro Area
230/2026	Gerhard Illing	Economic Theory and Central Bank Independence
229/2026	Sylvester Eijffinger, Jakob de Haan	Central Bank Independence: An Update
228/2025	Tobias Cwik, Ph.D., Christoph Winter	FX interventions as a form of unconventional monetary policy
227/2025	Prof. Michael D. Bauer, Ph.D., Miguel Acosta, Ph.D., Andrea Ajello, Ph.D.,	Financial Market Effects of FOMC Communication: Evidence from a New Event-Study Database

	Francesca Loria, Ph.D., Silvia Miranda-Agrippino, Ph.D.	
226/2025	Prof. Dr. Bernd Hayo, Dr. Johannes Zahner	Fiscal Talks: Parliamentary Debates and Government Expenditure
225/2025	Prof. Volker Wieland, Ph.D., Hendrik Hegemann	Moment of the Euro? Perceptions of US dollar Decline
224/2025	Prof. Volker Wieland, Ph.D., Hendrik Hegemann	ECB Policy and Strategy Review: Potential Improvements
223/2025	Prof. Volker Wieland, Ph.D.	Debt Sustainability Analysis: Assessing its Use in the EU's New Fiscal Rules
222/2025	Dominik Hecker, Maik H. Wolters	Nonlinear Estimation of a New Keynesian Model with Endogenous Inflation De-Anchoring
221/2025	Prof. Athanasios Orphanides, Ph.D.	Challenges for monetary policy and its communication
220/2025	Prof. Athanasios Orphanides, Ph.D.	Improving the ECB's policy strategy
219/2025	Ekaterina Shabalina, Mary Tzaawa-Krenzler	Heterogeneous Attention to Inflation and Monetary Policy
218/2025	Prof. Dr. Franz Seitz, Prof. Dr. Malte Krueger	Costs of Means of Payment for Consumers: Literature review and some sensitivity analyses
217/2025	Alexander Meyer-Gohde, Johannes Huber	Iterative Refinement of the QZ Decomposition for Solving Linear DSGE Models
216/2025	Michael Haliassos, Thomas Jansson, Yigitcan Karabulut	Wealth Inequality: Opportunity for Me or for Others?
215/2024	Michael D. Bauer, Eric Offner, Glenn D. Rudebusch	Green Stocks and Monetary Policy Shocks: Evidence from Europe
214 / 2024	Michael D. Bauer, Daniel Huber, Eric Offner, Marlene Renkel, Ole Wilms	Corporate Green Pledges
213 / 2024	Athanasios Orphanides	The Federal Reserve's Evolving Interpretation and Implementation of Its Mandate
212 / 2024	Matthias Rumpf Michael Haliassos Tetyana Kosyakova Thomas Otter	Do Financial Advisors Have Different Beliefs than Lay People?
211 / 2024	Michael Haliassos	Wealth Accumulation: The Role of Others
210 / 2024	Kamila Duraj Daniela Grunow Michael Haliassos Christine Laudenbach Stephan Siegel	Rethinking the Stock Market Participation Puzzle: A Qualitative Approach
209 / 2024	Balint Tatar Volker Wieland	Policy Rules and the Inflation Surge: The Case of the ECB
208 / 2024	Reimund Mink	Helmut Schlesinger: Wegbereiter und Garant der deutschen Geld- und

		Stabilitätspolitik wird 100
207 / 2024	Alexander Meyer-Gohde	Solving and analyzing DSGE models in the frequency domain
206 / 2024	Jochen Güntner Magnus Reif Maik Wolters	Sudden Stop: Supply and Demand Shocks in the German Natural Gas Market
205 / 2024	Alina Tänzer	Multivariate Macroeconomic Forecasting: From DSGE and BVAR to Artificial Neural Networks
204 / 2024	Alina Tänzer	The Effectiveness of Central Bank Purchases of long-term Treasury Securities:A Neural Network Approach
203 / 2024	Gerhard Rösl	A present value concept for measuring welfare
202 / 2024	Reimund Mink Karl-Heinz Tödter	Staatsverschuldung und Schuldenbremse
201 / 2024	Balint Tatar Volker Wieland	Taylor Rules and the Inflation Surge: The Case of the Fed
200 / 2024	Athanasios Orphanides	Enhancing resilience with natural growth targeting
199 / 2024	Thomas Jost Reimund Mink	Central Bank Losses and Commercial Bank Profits – Unexpected and Unfair?
198 / 2024	Lion Fischer Marc Steffen Rapp Johannes Zahner	Central banks sowing the seeds for a green financial sector? NGFS membership and market reactions
197 / 2023	Tiziana Assenza Alberto Cardaci Michael Haliassos	Consumption and Account Balances in Crises: Have We Neglected Cognitive Load?
196 / 2023	Tobias Berg Rainer Haselmann Thomas Kick Sebastian Schreiber	Unintended Consequences of QE: Real Estate Prices and Financial Stability
195 / 2023	Johannes Huber Alexander Meyer-Gohde Johanna Saecker	Solving Linear DSGE Models With Structure Preserving Doubling Methods
194 / 2023	Martin Baumgärtner Johannes Zahner	Whatever it takes to understand a central banker – Embedding their words using neural networks
193 / 2023	Alexander Meyer-Gohde	Numerical Stability Analysis of Linear DSGE Models – Backward Errors, Forward Errors and Condition Numbers

192 / 2023	Otmar Issing	On the importance of Central Bank Watchers
191 / 2023	Anh H. Le	Climate Change and Carbon Policy: A Story of Optimal Green Macroprudential and Capital Flow Management
190 / 2023	Athanasios Orphanides	The Forward Guidance Trap
189 / 2023	Alexander Meyer-Gohde Mary Tzaawa-Krenzler	Sticky information and the Taylor principle
188 / 2023	Daniel Stempel Johannes Zahner	Whose Inflation Rates Matter Most? A DSGE Model and Machine Learning Approach to Monetary Policy in the Euro Area
187 / 2023	Alexander Dück Anh H. Le	Transition Risk Uncertainty and Robust Optimal Monetary Policy
186 / 2023	Gerhard Rösl Franz Seitz	Uncertainty, Politics, and Crises: The Case for Cash
185 / 2023	Andrea Gubitz Karl-Heinz Tödter Gerhard Ziebarth	Zum Problem inflationsbedingter Liquiditätsrestriktionen bei der Immobilienfinanzierung